

普通高等教育"十一五"国家级规划教材

博弈论

Game Theory

□ 姚国庆 编著



高等教育出版社
HIGHER EDUCATION PRESS

VISIT…

LANZAROTE
Caliente.COM



普通高等教育"十一五"国家级规划教材

博弈论

Game Theory

□ 姚国庆　编著



高等教育出版社
HIGHER EDUCATION PRESS

图书在版编目(CIP)数据

博弈论/姚国庆编著.—北京:高等教育出版社,
2007.7

ISBN 978-7-04-021422-2

Ⅰ.博… Ⅱ.姚… Ⅲ.对策论-高等学校-教材
Ⅳ.0225

中国版本图书馆 CIP 数据核字(2007)第070805号

策划编辑 权利霞 **责任编辑** 顾 瑶 **封面设计** 王凌波
责任绘图 尹 莉 **版式设计** 王 莹 **责任校对** 金 辉
责任印制 韩 刚

出版发行	高等教育出版社	购书热线	010-58581118
社　　址	北京市西城区德外大街4号	免费咨询	800-810-0598
邮政编码	100011	网　　址	http://www.hep.edu.cn
总　　机	010-58581000		http://www.hep.com.cn
		网上订购	http://www.landraco.com
经　　销	蓝色畅想图书发行有限公司		http://www.landraco.com.cn
印　　刷	北京鑫丰华彩印有限公司	畅想教育	http://www.widedu.com
开　　本	787×960　1/16	版　　次	2007年7月第1版
印　　张	19.25	印　　次	2007年7月第1次印刷
字　　数	340 000	定　　价	27.00元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有　侵权必究

物料号　21422-00

序　言

传统经济学以研究资源的最优配置为核心，但仅仅如此是不够的，因为人是一种“社会性动物”，利益冲突是社会关系的本质特征，因而现代经济学与传统经济学的重要区别之一，就是将人类冲突与协调纳入经济学的研究范畴，试图对人类的经济活动给出更加科学的解释，这一努力的成果就是现在我们所看到的博弈论。博弈论不仅深远地影响了经济学的研究和发展，而且对其他社会科学，甚至是自然科学都产生了重大影响。博弈论的思想和方法已经成为现代经济学研究的基本思想和方法，离开了博弈论，现代经济学就名不副实。在其他社会和行为科学中，博弈论同样也得到了广泛运用。通过查阅近十年来的诺贝尔经济学奖获奖情况，我们就会得到更直观的认识——要获诺贝尔经济学奖，最好的捷径就是和博弈论挂钩。

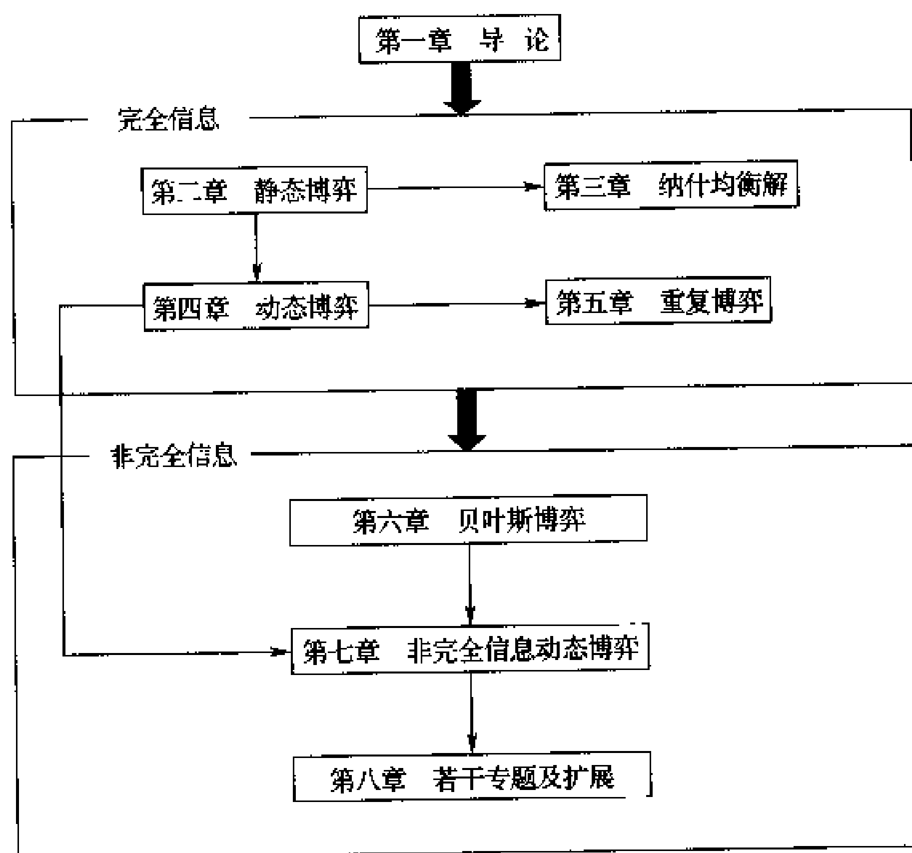
本书介绍的是博弈论的基本思想，力图全面但又重点突出，力求严谨又不失易懂，目的在于教会读者如何运用博弈论的基本思想去理解和解释我们所遇到的各种社会、经济、文化和政治现象。对于绝大多数人而言，学习博弈论的主要目的在于运用而不是研究。就像博弈论不仅仅是为经济学服务一样，虽然这本书主要是一本经济学教材，但是它也不仅仅是为学经济学的读者所写。

1．本书结构

本书的结构大体分为三个部分，第一部分为完全信息博弈，它由第一章至第五章组成；第二部分为非完全信息博弈，它由第六章至第七章组成；第三部分则介绍了博弈论理论的一些新发展，由第八章组成。它们之间的关系可用下面的图清楚地表示出来。图中细箭头表示前者是后者的基础，而粗箭头表示前一部分是后一部分的基础，也可以理解为随着学习的深入，模型变得越来越复杂，但复杂模型是由简单模型演化而来。由于本书主要是为大学经济学高年级本科生和研究生一年级学生学习博弈论之用，因而经济学的例子较多，但也尽量多地包含其他领域的例子，因为博弈论不仅仅是属于经济学的。为了突出运用，本书专门设置了三章，分别是第三章、第五章和第八章。第三章主要是介绍如何运用纳什均衡概念来求解完全信息静态博弈经典模型的均衡解，而第五章则是介绍如何运用子博弈完美均衡概念来求解完全信息动态博弈经典模型的均衡解，第八章则是介绍了博弈论最前沿的一些发展和运用。这样安排主要是基于如下几点考虑：一是运用是最好的学习方法，也是本书的目的之所在；二是打好基础是进

一步学习的根本;三是便于教师根据课时合理安排教学内容。从做学生的经历来看,很少有老师能把一本书全部讲完的,因而在教学安排上至少应有若干路径可供选择:

- 偏重理论的,可以重点讲解第二章、第四章、第六章和第七章。
- 强调基础和运用的,可以重点讲解第二章、第三章、第四章和第五章。
- 如果在微观经济学中已经学过完全信息博弈,那么可以重点讲解第六章、第七章和第八章。
- 如果时间较少,只是对博弈论作一介绍并培养学习兴趣,那么可以讲解第一章、第二章,并选择第三章中的一些经典模型进行讲解。
- 如果时间充裕,并且是经济学专业的,应讲解所有章节。



本书的篇章结构

无论采用哪种路径,都要求学生多做题,特别是书中的习题。对于博弈论来讲,如果不做题,学习效果就差了一大半。

2. 体例和说明

为了提高学习效果,本书的一个重要措施就是将习题完全融入每章的内容之中,而不是放在每章的最后,带思考的学习和在学中做是本书的一大特点。书中的习题分不同的难度,其难度用符号㊂、㊥和㊦表示,分别代表容易、中等和较

难。作为学生，最好能够独立完成容易和中等的习题，而较难的习题则可以通过小组讨论、教师点拨式的讲解等方式来完成。围绕习题来讲解博弈论的主要思想是一个非常好的学习方法和手段。

为了进一步巩固学生的学习效果，本书在内容上也作了相应安排，每章的开头是这一章内容的简略概括，最后则由关键词和知识要点组成。书中有些地方用楷体字书写，主要是一些补充证明和扩展，读者跳过并不影响阅读。此外，教材还补充了一些博弈论的课外知识，通常用方框框起以便与书中的主要内容相区别，相信这些内容也是值得一读的。书中的图和表，定义、定理和重要公式都是根据其在各章中出现的顺序来编排的，例如，图 3－1、定理 3.1 就表示第三章的第一个图，第三章出现的第一个定理。例题中出现的公式，除非必要，否则不编号，如果编号，为了与重要公式相区别，会在中间加上一个字母，例如4－P－1表示第四章中例题价格领先制中的第一个公式（P 为价格的头一个字母）。

3. 阅读对象

本书的主要阅读对象为大学经济学类高年级本科生或研究生一年级学生，他们在学习微观经济学时，已经接触到一些博弈论的基础知识，但并不深入，也不成体系。在本书中，作者努力做到对博弈论有一系统的交代，同时尽力做到深入浅出，重在使学生理解并能加以运用。为了能够很好地理解博弈论，要求学生具有一定的经济学知识和数学知识，但要求都不是很高，会求导数和掌握基本的概率论知识就能将本书读完。虽然作者会努力尽可能详尽地加以解释和说明，但思考还是读者自己的事。如果读者遇到不懂的地方，除了“好问”之外，多动笔也十分关键。“好记性不如烂笔头”，动笔自有一番新天地。最后，学好这门课还需要反复多看几遍本书，有机会的话阅读书后的参考文献就更好。本书虽然是一本经济学教材，但它同样也适合其他感兴趣的读者自学。

学习是一件痛并快乐着的事，如果你能读完本书，相信它能对你今后的人生有所助益。最后，我想用在网上看到的一道智力题来结束序言。关于这道智力题（实际上是一道博弈论题），网页上写到，如果你能在 20 分钟之内做出这道题，那么你就有资格和能力获得百万年薪。这道题是这样的：五个强盗抢到了一笔财富，然后进行分赃。首先由第一个强盗提出分配方案，如果多数人（包括半数）同意，那么财富分配结束，如果多数人不同意，那么就把第一个强盗杀掉，由第二个强盗提出分配方案；如果大多数人同意第二个强盗的方案，那么分配结束，否则把第二个强盗杀掉，再由第三个强盗提出分配方案。其他的以此类推，不再赘述。每一个强盗只有两种选择——同意或不同意，并且每个强盗在不损失利益的前提下总是避免杀人。如果你是第一个强盗，那你应该提出什么样的分配方案才能既把命保住又能分得最多？看完这道题后，我相信你一定有了最

大的动力来学习博弈论了(玩笑话)。

㊀ 习题1 如果你是第二个强盗,或第三个强盗,或第四个强盗,你的最优策略又是什么?这是一个博弈吗?如果是,它的均衡解是什么?

最后,祝大家一路顺风,到达最后一页!

4. 鸣谢

本书在作者以前工作的基础上进行了彻底的修改和补充,并在南开大学国家经济学基础人才培养基地进行了讲授,在这个过程中,得到了许多同学的无私帮助。令我出乎意料的是,其中有许多是学习历史、政治的学生,他们对博弈论知识的渴求与遭遇到的困难都给我留下了深刻的印象。如何使得博弈论更加亲民,而不是拒人于千里之外,就成为我写作本书的最大动力。在此,对帮助过作者写这本书的同学和老师表示深深的谢意!为了这本书的出版,南开大学经济学系和高等教育出版社做了大量辛勤工作,在此也深表谢意!

姚国庆

2007年2月

目　录

第一章　导论 …… 1

第一节　什么是博弈论 …… 1

第二节　博弈论的经济学渊源 …… 6

第三节　博弈论革命 …… 11

第四节　博弈论要点 …… 16

关键词 …… 24

内容要点 …… 24

进一步阅读 …… 25

第二章　完全信息静态博弈 …… 26

第一节　博弈的基本式 …… 27

第二节　纳什均衡 …… 33

第三节　简单运用举例 …… 42

第四节　最优反应函数 …… 50

第五节　对称博弈和对称均衡 …… 59

第六节　混合策略与纳什定理 …… 60

关键词 …… 76

内容要点 …… 77

进一步阅读 …… 77

第三章　纳什均衡运用举例 …… 78

第一节　伯川德模型 …… 78

第二节　多党竞选 …… 80

第三节　事故赔偿法 …… 83

第四节　公共地悲剧 …… 87

第五节　争议仲裁 …… 90

第六节　合作性谈判 …… 93

第七节　报案 …… 97

第八节　专家诊断 …… 99

第九节　连续行动下的混合策略纳什均衡 …… 103

关键词 …… 106






内容要点 …… 106
进一步阅读 …… 106
第四章 完全信息动态博弈 …… 107
第一节 完美信息与完全但不完美信息 …… 108
第二节 动态博弈的扩展式 …… 109
第三节 策略和结果 …… 112
第四节 纳什均衡与子博弈完美均衡 …… 115
第五节 经典举例 …… 122
关键词 …… 140
内容要点 …… 140
进一步阅读 …… 141
第五章 重复博弈 …… 142
第一节 重复博弈的定义及扩展式 …… 142
第二节 合作产生的原因 …… 145
第三节 无名氏定理 …… 156
第四节 经典举例 …… 162
关键词 …… 173
内容要点 …… 173
进一步阅读 …… 174
第六章 贝叶斯博弈 …… 175
第一节 豪尔绍尼转换 …… 175
第二节 贝叶斯博弈与纳什均衡 …… 180
第三节 经典举例 …… 189
第四节 显示原理 …… 212
关键词 …… 216
内容要点 …… 216
进一步阅读 …… 217
第七章 非完全信息动态博弈 …… 218
第一节 一般扩展式 …… 218
第二节 序列均衡 …… 223
第三节 经典举例 …… 234
第四节 序列均衡的再精炼与运用 …… 254
第五节 均衡概念小结 …… 266
关键词 …… 267

内容要点 …… 267
进一步阅读 …… 268
第八章 若干专题及扩展 …… 269
第一节 同盟博弈与核 …… 269
第二节 演化博弈与演化稳定策略 …… 273
第三节 学习博弈 …… 280
关键词 …… 283
内容要点 …… 284
进一步阅读 …… 284
参考文献 …… 285
人名索引 …… 288
主题索引 …… 289
教学支持说明

第一章

导 论

本章的目的在于使读者在学习博弈论之前对博弈论有一个总体上的认识。这种认识对于即将开始博弈论学习之旅的读者而言无疑是有益的,它就好比一张地图,为即将扬帆远洋的航船指明了方向和位置。考虑到本书的读者主要是在校的经济类本科生,他们已经接受了将近一年的经济学基础理论学习,但对博弈论却知之不多,有此导论无疑能起到承上启下的作用。

本章主要讨论四个问题,即何谓博弈论、博弈论的经济学渊源、博弈论革命及博弈论要点。其中,也希望澄清几个人们认识上比较模糊的地方,例如,长期以来,人们都认为博弈论实际上是数学的分支,这种认识是片面的。从思想根源来看,博弈论属于经济学的范畴更为适当,数学对于博弈论而言只是工具而已。同时,本章也希望通过对博弈论的经济学思想渊源的说明,为学习博弈论打下一个方法论基础。

第一节 什么是博弈论

在回答什么是博弈论之前,首先需要明白什么是博弈。按照《现代汉语词典》的解释,"博"是丰富多彩的意思,而"弈"则指下棋、打牌等对抗性的游戏,因而"博弈"就是指丰富多彩的对抗性游戏。在英语中,博弈一词最早出现于冯·诺伊曼和摩根斯坦恩所著《博弈与经济行为的理论》一书,其中博弈一词为英语"游戏(game)"一词的复数,表示各种各样的游戏,因此汉语中的"博弈"与英语中的"游戏(game)"意思完全一样。看来博弈与游戏结下不解之缘绝非偶然。下棋、打牌这类游戏的一个共同特征是所有游戏的参与者都试图"智胜"对手,要智胜对手就需要随机应变,最为重要的是能够预见到对手可能采取的行动,事先有所准备。用专业一点的话来说,就是游戏参与者彼此的策略是相互依存的。为了说明这一点,我们以石头、剪刀、布这个游戏为例。

【例1.1】 石头、剪刀、布

孙悟空和猪八戒护送唐僧来到火焰山,方圆百里酷热难当,孙悟空和猪八戒

二人必须有一人去寻找淡水,否则唐僧师徒将命丧火焰山。但唐僧师徒经过长途跋涉已非常劳累,在火焰山找水更是难上加难,因而寻找淡水是份苦差事,孙悟空和猪八戒谁都不想承当这份工作。唐僧命令谁去都会引来"凭什么让我去"这样的诘问。一个较好的解决办法就是让唐僧当裁判,孙悟空和猪八戒通过石头、剪刀、布来决定胜负,谁输谁去寻找淡水。由于双方赢的概率一样,因而无论谁输,都能输得心服口服(不存在耍赖这种情况)。游戏规则如下:孙悟空和猪八戒必须同时出招,谁后出算谁输;石头胜剪刀,剪刀胜布,布胜石头;如果出招相同,不算输也不算赢,重新再来一次。例如,孙悟空出剪刀,猪八戒出布,孙悟空赢。如果孙悟空和猪八戒都出剪刀,则重新开始游戏。根据上面所给出的信息,可以把孙悟空和猪八戒两人可能的策略和输赢列成矩阵形式,如图1-1所示。

		猪八戒		
		石头	剪刀	布
孙悟空	石头	未定,未定	休息,找水	找水,休息
	剪刀	找水,休息	未定,未定	休息,找水
	布	休息,找水	找水,休息	未定,未定

图1-1 石头、剪刀、布博弈矩阵

从这个游戏可以看出,无论是孙悟空还是猪八戒,他们出什么招,关键在于他们对对手可能出什么招的"猜测"上。如果孙悟空"猜测"猪八戒可能出石头,孙悟空的最优策略是出布,其他招都属不智;如果孙悟空"猜测"猪八戒可能出剪刀,孙悟空的最优策略是出石头,其他招都属不智;如果孙悟空"猜测"猪八戒可能出布,孙悟空的最优策略是出剪刀,其他招都属不智。同理,对于猪八戒而言也是一样。因而可以看到,孙悟空采取什么策略关键取决于猪八戒的策略,而猪八戒采取什么策略反过来又取决于孙悟空的策略。对这种情况正规的表述是:孙悟空和猪八戒的策略具有相互依存性。如果引入最优反应函数的概念,策略的**相互依存性**就表现得更为明显。假设 B_s 为孙悟空的最优反应函数,B_z 为猪八戒的最优反应函数;a_{11} 表示孙悟空的策略——石头,a_{12} = 剪刀,a_{13} = 布;a_{21} 表示猪八戒的策略——石头,a_{22} = 剪刀,a_{23} = 布。这里默认 a 的第一个下标表示的是第几个参与者,1 表示孙悟空,2 表示猪八戒;a 的第二个下标表示是第几个策略,1 表示石头,2 表示剪刀,3 表示布。由此可得,

孙悟空的反应函数:$B_s(a_{21}) = a_{13}$,$B_s(a_{22}) = a_{11}$,$B_s(a_{23}) = a_{12}$;

猪八戒的反应函数:$B_z(a_{11}) = a_{23}$,$B_z(a_{12}) = a_{21}$,$B_z(a_{13}) = a_{22}$。

即孙悟空的反应函数的自变量是猪八戒的策略,并由此得到孙悟空的最优策略,例如 B_s(石头) = 布,即如果猪八戒出石头,孙悟空的最优策略是出布。对于猪八戒而言,其反应函数的自变量则为孙悟空的策略,并据此得到自己的最优

策略。显然孙悟空和猪八戒的策略是相互依存的。

这种策略依存现象广泛存在于人类社会的方方面面,最极端的情景就是以命相搏的战争。这就是为什么博弈论思想的萌芽最早产生于战争,并最早运用于战争。从这点(策略依存)来看,战争、阶级冲突、政治斗争、经济竞争,甚至生物界的优胜劣汰都与游戏并无二致。这就是为什么博弈论不仅被运用于经济学的研究,而且被广泛运用于其他社会科学和自然科学。例如,博弈论被广泛运用于军事、政治、法律等社会领域,在自然科学中,则被运用于人工智能、物种演化等领域。为了加深理解,再举几个有趣的例子。

【例 1.2】 诺曼底登陆

诺曼底登陆是第二次世界大战中的重要战役,它对加速法西斯德国的垮台具有重要的战略意义。当时盟军登陆的地点有两个,一是加来,二是诺曼底。由于盟军的军力有限,不可能同时在两个地点登陆。由于相同的原因,德军也不可能同时在两地设防,只能集中兵力在一地设防。如果盟军在加来登陆,而德军正好在加来设防,那么由于德军以逸待劳,再加上天时地利,将能够击败盟军的进攻。相反,如果盟军在加来登陆,而德军在诺曼底设防,盟军将不会遇到有效的抵抗,能够登陆成功。其他同理。可以将历史上的诺曼底登陆抽象为图 1-2。

		德军	
		加来设防	诺曼底设防
盟军	加来登陆	失败,成功	成功,失败
	诺曼底登陆	成功,失败	失败,成功

图 1-2 诺曼底登陆博弈矩阵

从图 1-2 可以看出,盟军之所以在诺曼底登陆,是因为盟军相信德军会在加来设防;而德军之所以在加来设防,是因为它相信盟军会在加来登陆。用学术一点的话来说,就是盟军的最优策略是德军策略的函数;反之,德军的最优策略是盟军策略的函数,即盟军和德军的策略是相互依存的。

【例 1.3】 鸽派和鹰派

苏联和美国是冷战时代的两个超级大国,并长期处于对抗状态。从历史来看,苏联和美国很少出现公开的直接冲突,通常是美国强硬,苏联就会采取暂时回避的策略,如果是苏联强硬,美国就会主动回避。像古巴导弹危机那样的紧张局面极少出现,即使出现,双方都会采取妥协的策略。苏联和美国之间的这种现象可以用一个简单的博弈模型来加以概括。显然,如果苏联和美国之间发生直接冲突,后果有可能是同归于尽,因而妥协比一味强硬结果要好。在政治术语中,鸽派通常代表妥协路线,而鹰派通常代表强硬路线。苏联和美国之间的博弈如图 1-3 所示。

		美国	
		鸽派政策	鹰派政策
苏联	鸽派政策	0,0	-1,+1
	鹰派政策	+1,-1	-∞,-∞

图 1-3 鸽派和鹰派博弈矩阵

图 1-3 表明,如果苏联和美国都采取鸽派政策(避免冲突),那么双方得到的利益为零,表示双方相安无事。如果双方都采取强硬路线,那么必然导致冲突的发生,后果是双方同归于尽,得到的利益为负的无穷大。如果一方采取强硬路线,而另一方采取鸽派政策,那么实行鹰派政策的一方可以威吓对方从而获得较多的利益,用 +1 代表,而实行鸽派政策的一方则失去一部分利益,用 -1 代表。因而,苏联和美国到底采取什么政策,主要就取决于彼此认为对方可能采取的策略。如果美国采取鸽派政策,那么苏联的最优反应就是表现出强硬立场。如果美国采取鹰派政策,那么苏联的最优策略就是退避三舍,采取鸽派政策;反之亦然,上述这个模型虽然简单,但却真实地反映了冷战时期的本质特征。在图 1-3 中,存在着两个博弈均衡解,它们分别是(苏联鸽派政策,美国鹰派政策)和(苏联鹰派政策,美国鸽派政策)。这就为冷战时期美苏两大超级大国主动避免冲突的行为提供了有说服力的解释。

㊥ 习题 1.1　请举出若干博弈例子,并说明博弈参与者的策略和收益。

通过上面的说明,现在给出博弈的正式定义。

定义 1.1　**博弈**是指利益存在冲突的决策主体(个人、企业、集团、政党、国家等)在相互对抗(或合作)中,对抗双方(或多方)相互依存的一系列策略和行动的过程集合。

在这个定义中有几点要注意:

第一,博弈中的参与者各自追求的利益具有冲突性。如果决策主体之间的利益是一致的,就不是博弈。从学术观点来看,即使一个博弈包含无穷多参与者,但利益一致,就可以理解为一个人。显然一个人是不会和自己博弈的。因而从某种意义上看,博弈论是一门研究冲突的学科,它为人们理解冲突和合作提供了一种重要的思想方法。

第二,博弈是一个过程集合。博弈不是一个孤立的事件,而是人们在对抗过程中有关的所有方面的集合。它包含参与者的集合、策略的集合、行动的集合、信息的集合等。把博弈看作一个集合是思维从具体到抽象的重要一步。

第三,博弈的一个本质特征就是策略的相互依存性。如果博弈参与者之间的策略不存在依存性,那么与一个人自娱自乐的游戏并无区别,当然也就不能称其为博弈。不过在一种特殊的境况下,有一种博弈"不存在"策略的相互依存性,这种博弈就是包含严格优策略的博弈,例如图 1-4 所示的囚徒困境。

		囚徒2	
		沉默	招供
囚徒1	沉默	-1，-1	-9，0
	招供	0，-9	-6，-6

图1-4 囚徒困境博弈矩阵

【例1.4】 囚徒困境

警方抓住了两个犯罪嫌疑人——囚徒1和囚徒2，但是警方掌握的证据并不充分，因而需要犯罪嫌疑人的口供。警方将两个囚徒分别关在不同的房间，但告诉两个囚徒相同的信息，即如果一方沉默而另一方招供，那么招供方立即释放，沉默方将被判入狱9个月；如果双方都沉默，那么两人将在一个月后因证据不足而被释放；如果双方都招供，那么两人都将被判入狱6个月。在这种情景下，对于任何一个囚徒而言，无论对方选择什么策略，他的最优策略都是招供，因而出现了囚徒1"不依赖"囚徒2的策略来选择最优策略的情况（反过来也一样）。招供就是两个囚徒的严格优策略。

博弈论简史

博弈论的思想最早可以追溯到18世纪，但真正作为一种理论研究则始于20世纪20年代，公认的开山之作是1944年出版的由科学家冯·诺伊曼和经济学家奥斯卡·摩根斯坦恩合著的《博弈与经济行为的理论》一书。冯·诺伊曼和奥斯卡·摩根斯坦恩认为，以往的数学是为物理学的发展而建立起来的，它描述的是客观世界行为，对经济学却不太合适。他们认为经济学所研究的对象更像是一场游戏中的参与者，相互之间预期对方的行动；因此，描述、观察研究对象就需要一系列新的数学工具。这一套新的数学工具，就被他们命名为博弈论。"博弈论"这个名词虽然提出来了，但却找不到发展的突破方向。直到1951年纳什提出了"纳什均衡"的概念，博弈论才找到正确的方向，并取得了长足的发展，开始运用于经济学、政治科学和心理学。到了20世纪70年代，博弈论又被用于生物学和计算机科学，并逐渐形成一个完整的体系，在经济学中也逐渐取得主流地位。1994年经济学诺贝尔奖授予了三位对博弈论发展作出贡献的学者，确认了博弈论对人类思想的贡献。

明确了什么是博弈，对博弈论就好下定义了。

定义 1.2 博弈论是专门研究博弈如何出现均衡的规律的学科。

正是由于均衡在博弈论中具有中心地位（均衡在经济学中也同样如此），因而博弈在经济学中运用最广，并主要由经济学家加以发展。1994 年诺贝尔经济学奖授予了对博弈论发展作出贡献的三位学者，其中一位是数学家，两位是经济学家。另外，正是由于均衡的重要，我们将 Game Theory 译作"博弈论"而不是"对策论"，因为对策论是一个数学命题，而均衡是一个经济学命题。按照《中国大百科全书》的解释，均衡是"经济学中，用以分析考察经济中若干变量之间的关系，以解释经济现象及其变化的概念或分析方法"。因而博弈论体现的是一种经济学思想，不过这种思想不仅仅局限于经济学本身。

博弈论将重点放在博弈的均衡上，有两点好处：一是使得博弈论具有广泛的适用性，适用性越强，理论价值越高；二是使得博弈论研究有了立足的基点，不至于迷失方向。

㊂ 习题 1.2　请尽可能多地列举出博弈论可以运用的领域。

实际上，只要存在利益冲突的双方有可能形成某种能够得到维持的状态的情形，就是博弈论可以一展所长的地方，例如，商业竞争、党派选举、外交活动、军事冲突，等等。

第二节　博弈论的经济学渊源

通过博弈论的定义，我们已经知道博弈论与经济学存在着不解之缘，接下来将进一步明确这种关系，并对博弈论的方法论基础加以说明。

一、模型

什么是模型？通俗地说，模型就是一个逻辑体系。它从人们的实践活动中抽象而来并帮助人们理解所观察到的现象。马克思主义哲学中的从具体到抽象，从抽象到具体就是这个意思。博弈论实质上就是由若干个模型所组成。作为抽象，模型离不开简约化，其理论价值也恰恰在于简约化。一幅"真实"的地图是没有任何理论意义的，地图的价值就在于它是简略的。一个理论模型通常由三段组成：理论假设（或理论前提）——→逻辑推理——→理论假说（或理论结论）。模型的简约化通常发生在理论前提上，简约化到什么程度关键看模型试图解答的问题。例如，当我们计算北京天安门到北京火车站的距离时，可以假设地球是平面的，忽略地球实际上是椭圆的；但是当我们要计算从北京到纽约的距离时，椭圆的地球就是关键的前提而不能省略。因而我们看到，模型的前提是否

真实并不是判断一个模型“好坏”的标准。模型的简约化实际上是一把双刃剑，它不仅是科学也是艺术，甚至有时运气也很关键，弄不好会使得模型得出完全错误的结论和认识。至于逻辑推理，由于逻辑学、数学等的发展，发现错误相对容易。而对于理论假说，则要求假说必然是对某种现象关系的阐述，并能在实践中观察到，从而与实际中的现象对照，判断其是真是伪。如果真，理论得到“某种程度的支持”；如果伪，检查是假设错，还是推理错，从而进行理论修正，或者废弃理论。

为了更好地现解模型的意义，不妨举一个著名的例子。学过世界历史的人都知道，早在古希腊时代，人们就知道地球是圆的。有人可能会觉得奇怪，古希腊的人是怎么知道地球是圆的呢？这只不过是一个理论模型得出的结论，见图1－5。

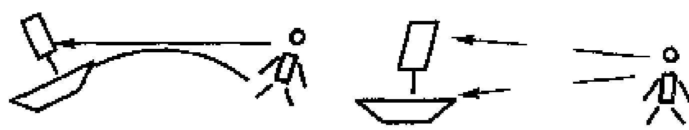


图1－5　地球是圆的，还是平的？一个理论模型

这个模型是这样的：

(1) 假设：地球是圆的。

已知：光线是按直线行走的。(现实中观察到的事实。)

那么：在港口的人，首先看到远方驶来船只的船桅，并逐渐看到船的下部。(与现实中观察到的事实相符。)

(2) 假设：地球是平的。

已知：光线是按直线行走的。(现实中观察到的事实。)

那么：在港口的人，首先看到的是整个船身。(与现实中观察到的事实相悖。)

结论：地球是圆的。

做出上面推断的人就是大名鼎鼎的希腊智者——亚里士多德。

下面再举个具体的例子来看看。前面我们曾提到过囚徒困境。囚徒困境可以说是博弈论里面最著名的模型，因为博弈论的发展曾受到囚徒困境的启发。可以毫不夸张地说，博弈论的主要思想都可以用囚徒困境及其扩展来概括。

模型——囚徒困境

假设：两个囚徒都是理性的(即都追求自身利益最大化)；

已知：两个囚徒被抓，不存在串谋，坦白从宽、抗拒从严；

结论：两个囚徒都会招供。

在实际审讯犯罪嫌疑人时，坦白从宽、抗拒从严虽然不是百分之百的管用，

但在绝大多数情况下确实发挥关键作用。在欧美国家，坦白从宽、抗拒从严这种制度同样是存在的，就是所谓的污点证人。

实际上囚徒困境只要稍作变动，就可以用来说明商业竞争。例如在我国，前几年手机还是身份的象征，而现在卖菜的老大娘都有，它已经变成普通的消费品。相反，电信服务却要差很多，长期以来双向收费、服务价格昂贵，出现买得起手机，用不起手机的现象。这是为什么？通过类似于囚徒困境的模型就可以回答这个问题。

模型——手机市场和电信市场竞争

假设：手机生产商和电信服务商都是理性的（即追求利润最大化）；

已知：手机生产商较多，电信服务商较少，消费者总是选择价廉物美的产品；

结论：手机生产商的竞争要激烈得多，只有不断推陈出新、降低价格才能赢得更多消费者；而电信服务商较少，存在较强垄断，因而不用提高服务质量，降低价格就能获得较多消费者。这类似于博弈论中所说的“一个竞争者胜过十个监督者”的情形，所以就出现了买得起手机，用不起手机的现象（后面章节还会对它进行具体的分析）。

最后，需要记住的是，当理论假说与实际相符时，我们只是说“理论得到某种程度的支持”，并没有说理论得到了证明，在地球是圆的还是平的理论模型中，真正的证明要等到哥伦布的航行完成之时。而有时候，直接的证明根本是不可能的。

㊥ 习题1.3 什么是模型？你能根据自身的体验提出一个模型来解释身边发生的经济现象吗？

二、博弈论的经济学思想渊源

前面反复提到一个重要的词，即博弈的参与者是“理性”的。那么什么才是理性？在标准的经济学中，所谓理性是指决策主体的偏好是理性的。而经济学中的理性选择理论就构成了博弈论的一个思想基点。

可以把参加博弈的人所能选择的行动和策略看作是一个集合（回忆博弈的定义），称为选择集，定义在这个选择集上的排序就是所谓的**偏好**。从理性的角度出发，博弈的参与者总是希望在各种约束和规则的限制下，尽可能地实现自身利益的最大化，因而定义在选择集上的偏好就意味着理性的参与者总是可以找到最优的行动和策略组合。如果这个结果对于参与者各方而言都是最优的，那么这种状况就会得到维持，参与者各方都没有动机去改变这种状况，即形成了所谓的均衡。那么是不是只要参与者是理性的，就一定会出现最优的选择？答案是肯定的。在微观经济学中这被称为“理性选择原理”。

理性选择原理：如果决策主体的偏好是理性的，那么（有限）选择集中就一定存在最优选择，这个选择可能是唯一的，也可能是多个。

理性选择原理不仅构成了博弈论的思想出发点，实际上也是经济学得以建立的根基。

㊐ 习题1.4　什么是理性选择原理？其中包含哪些要点？可以参考微观经济学的内容来回答这个问题。

那么什么样的偏好才能算是理性的？在严格偏好关系$\succ$下，如果满足下面两个条件，那么偏好就是**理性**的，否则就是非理性的[①]。

(1) 偏好是非对称的，即任意来自选择集的两个选择x和y，$x \succ y$和$y \succ x$不能同时成立。

(2) 偏好是负传递的，即对于来自选择集的选择x,y和z，如果$x \succ y$，那么要么$x \succ z$，要么$z \succ y$或是两者都成立。

在具体表述偏好时，经济学中使用效用函数这个工具，而在博弈论中，使用的却是收益函数(payoff function)。虽然名字不同但实质一样。如果用$u(\cdot)$来表示博弈中的收益函数，用a来表示参与者的行动，那么：

当且仅当博弈的参与者偏好a_1胜于a_2且有$u(a_1) > u(a_2)$时，u是博弈参与者的收益函数。

由于收益函数表示的是一种排序关系，因而表示某种偏好关系的收益函数并不唯一，这在微观经济学中被称为效用函数的**单调变换**。

㊐ 习题1.5　参与者有三个行动a_1,a_2,a_3，其偏好关系为$a_1 \succ a_2$，$a_2 \succ a_3$，有$u(a_1)=1\,000$，$u(a_2)=1$，$u(a_3)=0.9$和$v(a_1)=10$，$v(a_2)=0.01$，$v(a_3)=0.009$，那么收益函数u和v表示的偏好是否一样？还存在其他的收益函数吗？

实际上，收益函数表示的只是一种序数关系，因而收益函数之间的数值差额的多少并不存在经济学上的意义。

有一点需要略加说明。通常人们习惯把理性等同于自私自利，这有一定的道理，因为追求自身利益的最大化通常会导致自私行为。但需要指出的是，理性并不等同于自私自利，它同样可以包含利他主义。例如，雷锋就是一个理性的人，因为从雷锋的日记和所作所为来看，怎么看雷锋都不是一个非理性的人，他是那个时代的英雄和成功者，但雷锋却是一个利他主义者。只不过在雷锋的偏好中，不仅包含着自己的幸福，而且包含着他人的幸福。从偏好理性的定义来看，雷锋无论如何都是一个理性人，绝对不是非理性。因而理性是一个外延比自私自利大得多的概念，把两者等同显然是错误的。

---

① 根据严格偏好关系可以相应地定义弱偏好关系$\succeq$和无差异关系$\sim$。

(易) 习题1.6 某人既关心他的收入,也关心另一个人的收入。确切地说,就是他将另一个人的1.5单位收入等同于自身1单位的收入。例如,他认为他的收入为1,另一个人的收入为0,即组合(1,0),与他的收入为0,另一个人的收入为1.5,即组合(0,1.5),是无差异的。那么他面对组合(2,4),(0,10)和(3,0),他的偏好是什么?并给出一个与该偏好相一致的收益函数。他是一个利他主义者吗?

理性选择理论的局限性

任何理论都存在局限性,理性选择理论也同样如此。理性选择理论的局限性体现在两个方面:一是当人们在不确定的环境下作决策时,常常与理性偏好的要求相违背,产生悖论。例如,当只是改变问题的表述方式而问题本身未变时,人们的选择会相应发生改变。2002年经济学诺贝尔奖获得者之一的卡尼曼(Kahneman)教授的主要贡献就是对人们的理性预期行为进行了深入研究,得到了一些重要的研究成果。其研究领域成了一门新的学科,即行为经济学。二是如果将理性偏好推向极致,有可能导致非常荒谬的结论。所谓的"旅行者困境"就清楚地说明了这一点。两个旅行者在旅游胜地买了两个花瓶,但在回程途中被航空公司打碎。航空公司知道花瓶大约值100元,但并不清楚花瓶的确切价格。于是,航空公司要求两位旅客各自写下花瓶的价格,并按照两个旅客中所写的最低价格进行赔偿(航空公司认为写最低价格的旅客讲的是真话)。为了鼓励旅客讲真话,规定对讲真话的旅客奖励2元,对讲假话的旅客罚款2元。容易证明,在理性偏好的假设下,这个游戏唯一的结果是两人都写0。"旅行者困境"形象地说明,当我们运用理性偏好时,必须有所顾虑,而不能将其推向极致。任何理论都有其局限性,当问题超过了理论适用的范围时,理论就会失效。

人类社会不仅存在冲突,而且还存在广泛的合作,没有人与人之间的合作,人类社会早就荡然无存了。如果将理性人仅仅局限于自私自利,显然理论的解释力就要大打折扣。在博弈论中,合作博弈与非合作博弈同等重要,由于篇幅所限,本书主要涉及非合作博弈的内容。

通过上面的介绍，我们已看到，博弈论本身的发展离不开经济学，它从产生到发展都得益于经济学。那么它对经济学的发展又起到了什么样的作用呢？

第三节　博弈论革命

在经济学界，还没有人把博弈论的发展对经济学的影响称为革命，但客观看来，博弈论带给经济学的东西不仅仅是一个“革命”所能概括的，尽管这种影响是在一种潜移默化并且漫长的过程中完成的。它平平淡淡，但远胜轰轰烈烈。自20世纪90年代以来，近十年的时间里有4次诺贝尔经济学奖授予了与博弈论相关的研究领域（1994年的博弈论，1996年和2001年的信息经济学，2005年的冲突与合作），总共有10位经济学家因此而荣获诺贝尔经济学奖，是获奖人数最多的经济学研究领域（排第二位的是金融领域）。窥一斑而知全豹，再过十年，博弈论对经济学的影响会表现得更为明显。

经济学研究离不开四个重要范畴：决策主体、行为、制度结构和均衡。

决策主体，又称经济行为主体（actor）。经济行为主体是经济生活中追求自身利益最大化的决策者和参与者。例如消费者、生产者、政府以及任何的利益集团，都可以在经济学中被看作是决策主体。

行为（behavior）。行为是经济行为主体在面对各种约束条件下的最大化行动。最典型的例子就是消费者在收入、市场价格等的约束下使自身的效用最大化；生产者在成本、技术、市场价格等的约束下使利润最大化。

制度结构（institutional structnre）。制度结构可以看作是经济行为主体面对的所有约束，它规定了经济行为主体拥有什么条件去选择、如何去选择、能选择些什么，同时它也由行为主体自身的选择和其他行为主体的选择以及彼此互动关系所决定。

均衡（equilibrium）。均衡就是一个能够得以维持的结果，或者说是所有行为主体不得不接受（可能愿意也可能不愿意）而又不可能更好的结果。

以上四个范畴构成了经济学研究的基石。然面在传统的经济学研究下，通常主要考察的是行为主体如何选择，最终如何得到均衡，面把制度结构作为给定的外部条件来加以处理。改变不同的制度结构（实际上在经济学中这相当于更改约束），就得到不同的均衡结果。但是对于制度结构本身的性质和演变，传统经济学既不将它作为考察的对象，也不具有这样的能力。然而，排除对制度结构的考察对经济学而言是一个巨大的遗憾，这不仅大大制约了经济学的研究领域，而且经济学的有效性和科学性也受到了广泛的质疑。因面考察制度结构可以说是现代经济学和传统经济学的一个根本区别。

一般均衡理论是整个经济学的理论基石和道义基础。一般均衡理论表明市场机制是完美的。这一点集中体现在福利经济学第一定理和第二定理上。在完全竞争的市场和民主政府下，平等与效率可以兼顾。然而在下列情况下，一般均衡理论是不成立的。

（1）非完全竞争。在现实的经济生活中，特别是进入20世纪后，垄断越来越成为一种普遍而占主体的经济现象和制度结构。在垄断下，完全竞争的效率不复存在。

（2）外在性。由于行为主体之间客观存在着密切的联系，因而个体并非是孤立地存在着，一个行为主体的行为有可能对其他人造成影响，这种现象就是外在性，例如污染。外在性的存在表明市场不可能把所有的成本收益都计算在内，有些成本和收益客观地存在着，但市场却无能力对此进行计算。这说明市场是有缺陷的。

（3）公共产品。在现实生活中，公共产品无处不在，可以说它和私有物品同样普遍。但是在一般均衡理论中，公共产品是不考虑的。其原因在于对公共产品的消费存在着“搭便车”的行为，而市场机制本身无法解决“搭便车”问题。

（4）逆向选择问题。随着经济学研究的深入，人们发现信息不对称同样会造成市场的失灵，其中最突出的是**逆向选择**和**道德风险**问题。2001年诺贝尔经济学奖的三位得主就是因在这两个领域做出了开创性的工作而获奖。阿克洛夫（Akerlof）提出的“柠檬”（lemon）问题清楚地表明市场不仅不能解决这类问题，相反在逆向选择下会彻底地消失掉。针对这种情况，斯彭思（Spence）提出可以通过向行为主体发信号来解决逆向选择问题。现在，信号博弈已发展成为信息经济学的两大模型之一。

（5）道德风险问题。科斯（Coase）的企业理论被认为是新制度经济学的奠基之作。科斯认为之所以产生企业，是由于市场与企业（制度）相比交易费用太高，因而出现了企业，并替代市场。但是对于为什么市场交易费用高、交易费用的性质是什么等这些根本性的问题，科斯的回答却是不能令人满意的。在随后的研究中，人们才逐步发现这些问题不过是道德风险问题的一部分。因而如何解决道德风险成了新制度经济学和企业理论的核心课题，并在此基础上提出了各种各样的委托—代理模型。

对于上面所谈的市场失灵的五个方面，传统的经济学缺乏有效的手段去加以研究。由于一般均衡是在排除这五个方面后得出的结论，因而当面临这五个问题时，一般均衡就有可能是不存在的。这成了经济学的致命伤。一个世纪以来，经济学的主要发展实际上都是围绕着这五个问题展开的，例如垄断竞争理论、产业组织理论、企业理论、信息经济学、新制度经济学、不确定条件下的决策

（投资理论）以及宏观经济学。

博弈论的发展与成熟实际上使得经济学在面对上述五个问题时，有了进行深入研究的基础，而逆向选择和道德风险问题本身就是博弈论在发展过程中提出的。此外，博弈论使得经济学研究制度结构的来龙去脉及其演变成为可能。对于传统经济学而言，制度结构是作为外生变量给定的，因而经济学实际上几乎成了专门研究给定条件下如何规划利益最大化问题的学科，而这实际上是数学就能胜任的事。经济学日益丧失它的社会性，必然导致经济学的贫乏和失去其存在的必要。因而从这个意义上看，博弈论对于经济学而言是一场有着深远影响的革命。它极大地丰富了经济学的内容，使得经济学的研究领域和适用范围大大扩展。也使得经济学家的思维从一种静态、孤立的模式发展成互动、联系的模式，从而彻底改变了经济学家在观察问题和解决问题时的思维方式。这一点连马克思主义学者都加以承认。由于马克思关心现代社会的矛盾及其解决，因此博弈论简直就是为马克思主义的进一步发展所“定做”（tailor - made）的工具[①]。

㊥ 习题1.7　博弈论对经济学具有什么样的意义？你同意博弈论革命这种说法吗？理由是什么？

博弈论革命不仅表现在经济学中，也体现在其他社会科学中。

萨缪尔森（Samuelson）在其着作《经济学》中曾指出，社会科学研究中人们常常会犯“合成谬误”（fallacy of composition）式的推理错误，即认为个体理性必然意味着集体理性。这一推理的错误性可以通过下面的例子看出。当许多人坐着看表演时，坐在后面的人站起来可以看得更为清楚，因面站起来对个人而言是一种理性行为。但当所有人都站起来后，除了第一排的人看得清楚以外，其他人的境况都会变得更差。又如股市上个人抛出股票实属正常，但当所有人都抛出股票时，面对的可能就是股市崩溃。对于个人而言，跟随大流显然是理性行为，但个体理性导致的却是整体非理性。前面提到的囚徒困境，更反映出个体理性与集体理性是冲突的（个体理性要求都招供，面集体理性要求都沉默）。

正是基于这些观察，美国著名的政治学家奥尔森（Mancur Olson）于1965年发表了经典性著作《集体行动的逻辑》[②]，探讨了利益集团集体行动的微观基础。在奥尔森看来，利益集团的集体行动的目标是给每个成员提供公共产品。但公共产品具有非独占性和非排他性，因而存在着**搭便车行为**。这意味着在利益集团内部，集体行动的目标和个体理性存在着矛盾，集团并不像一个人那样有理

① Elster J. "Marxism, Functionalism and Game Theory". *Theory and Society* 11, 1982: 453 ~482

② 奥尔森．集体行动的逻辑．上海：上海三联书店，1990

性，而集体行动能否成功实际上又取决于集团内每一个成员的努力。因而奥尔森从单个成员对集团的公共产品的贡献与他从集体产品中得到的份额的关系入手，以此来对不同集团的内部结构进行分类。下面是奥尔森描述“集体行动的逻辑”的最简单的方程

$$P_i = R_i - C$$

其中 C 是获取公共产品的总成本，R_i是第 i 成员从总的公共产品中得到的总收益，P_i是第 i 成员的净收益。若对某些（少数）i 而言，$P_i > 0$，则奥尔森定义这样的集团为“含有特殊人物的集团”，若对所有 i 而言，$P_i < 0$，则奥尔森定义这类集团为隐集团。

在“含有特殊人物的集团”中，这些特殊人物从公共产品中得到的收益或好处是如此之大，以至于超过获取公共产品的总成本，因此他们愿意负担全部成本以获得公共产品。如果一个集团中含有这样的特殊人物，那么该集团在获得公共产品方面就能够成功。例如，在股份公司中，大股东就是利益集团（全体股东）中的特殊人物，通过大股东对经理实行监督，小股东“搭便车”得到丰厚的回报。例 1.5——富人穷人博弈也说明了这种情况。

【例 1.5】 富人穷人博弈

在一个小区内，住着一个富人和一个穷人。组织夜间巡逻能够有效防止偷盗，但夜间巡逻的成本为 4。假设富人的财产为 8，穷人的财产为 2。如果两人都巡逻，那么巡逻成本由两人均摊；如果只有一人巡逻，则由巡逻者承当。如果富人和穷人都巡逻，那么富人的收益为 8 - 2 = 6（财产 - 成本 = 收益），而穷人的收益为 2 - 2 = 0；如果富人巡逻而穷人不巡逻，那么富人的收益为 8 - 4 = 4，穷人“搭便车”，财产得保，收益为 2 - 0 = 2；如果两人都不巡逻，财产被窃贼偷光，收益都为零。从图 1 - 6 可以看出，如果富人巡逻，那么穷人最优的策略是不巡逻，“搭富人的便车”。问题是，在穷人不巡逻的情况下，富人会不会巡逻。答案是肯定的，因为巡逻有 4 的收益，不巡逻收益为 0。因此该博弈的均衡解是（富人巡逻，穷人不巡逻）。富人就是利益集团（住宅小区居民）中的特殊人物。

		穷人	
		巡逻	不巡逻
富人	巡逻	6,0	4,2
	不巡逻	8,0	0,0

图 1 - 6 富人穷人博弈矩阵

在“隐集团”中，奥尔森又区分出“大集团”和“中等集团”。此时，由于 $P_i < 0$，没有特别人物自愿承担公共产品的全部成本，每个成员从个人理性出发，

将不会自愿分担公共产品的总成本,原因是他们自以为:

(1) 由于集体产品的非排他性,不付出成本也可以享受;

(2) 自己付出努力,但并不能保证别人也会分担公共产品的成本,即存在着“搭便车”行为;

(3) 每个人对公共产品的贡献都很小,无关大局。

据此,奥尔森得出了他的“集体行动的逻辑”的主要结论:小集团要比大集团有力量,因为搭便车的现象要少。对于“隐集团”来说,如果没有外部强制或对成员的特别激励,将不会有提供公共产品的集体行动。

通过上面的介绍可以看到,奥尔森所谓的集体行动的逻辑实际上与囚徒困境的逻辑完全一样。1971 年,哈丁(Hardin)在其论文中,证明了“集体行动的逻辑”与 n 人囚徒困境实际上是等价的[①]。

罗尔斯(John Rawls)的《正义论》是 20 世纪下半叶人类思想史上的一部重要著作。自 1971 年出版以来,受到了伦理学家、政治学家、经济学家、社会学家和法学家的极大关注,此书的影响深度在 20 世纪下半叶难有出其右者。

在《正义论》中,罗尔斯试图证明在“原始状态”的“无知之幕”的条件下,人们如何通过理性的选择得到正义原则,以在一组“社会的基本结构”上配置各个社会成员的“基本益处”。“原始状态”下的“无知之幕”是指每一个人都不知道自己的出身、社会地位、财富、能力、智慧和爱好,从而保证没有人能够在选择和设计正义原则时塞进“私货”。“基本益处”在罗尔斯看来是每一个人都想要的东西,它包括基本的自由、机会、收入、财富和权力,以及(罗尔斯认为特别重要的)自尊。要配置“基本益处”需要一系列社会的、政治的和法律的、经济的制度。这个制度集合就是社会的基本结构,例如,宪法、所有制、市场机制、一夫一妻制等。罗尔斯在推导正义原则时,假设每一个人合理的选择是遵循“极小极大”的决策原则,在“极小极大”策略下,人们将发现罗尔斯所提出的正义原则要优于其他的正义原则(例如功利主义原则)。

从根本上说,社会契约是非零和博弈的合作解[②],而罗尔斯旨在“将洛克、卢梭和康德所代表的传统社会契约论一般化和达到更高度的抽象”[③],因此罗尔斯的分析实际上完全等同于博弈论中的动态博弈。又由于“极小极大解”只不过是博弈论中最原始的解概念之一,因面罗尔斯的分析又是远远不够的。博弈论

---

① Hardin R. "Collective Action as an Agreeable n - Prisoner's Dilemma". *Behavioral Science*, 1971, 5: 472 - 479

② Taylor M. *Anarchy and Cooperation*. London: Wiley, 1976

③ Rawls J. *A Theory of Justice*, VIII. Cambridge: Harvard University Press, 1971

的发展不仅指出了罗尔斯《正义论》的不足，同时也指明了继续前进的方向。

㊥ 习题1.8 囚徒困境中的极小极大策略解是什么？

在自然科学领域，博弈论同样有着广阔的运用前景，在人工智能、生物种群的演化等领域都显示出巨大的威力。1997年计算机"深蓝"战胜被称为棋王的国际象棋大师加里·卡斯帕罗夫，曾引起世人的巨大关注。深蓝的胜利与其说是机器的胜利，不如说是人类思想的胜利，因为"深蓝"下象棋的程序就是按照博弈论中的动态博弈的思想来编写的。

综上所述，把博弈论的发展对人类思想的贡献称之为"革命"并不为过，至少对经济学是如此。

第四节 博弈论要点

本节对博弈作一个总体但是简单的介绍，正式的叙述会在专门的章节中展开。

一、博弈的四个要素

博弈的形式多种多样，理论的力量就在于它能从各种具体的形式中抽象出一般。通过前面的例子可以总结出，任何一个博弈必然包含四个要素，这四个要素也是从理论上描述一个博弈不可或缺的材料。

(1) **参与者(player)**。参与者是参与博弈的决策主体。判断博弈参与者的根本标志是他是否是博弈的利害关系者，只有在博弈中存在利害关系的决策主体才被看作是博弈的参与者。一个较为特殊的、与其他参与者不存在利害关系的参与者定义为自然(nature)[①]。

(2) **博弈规则(rule)**。博弈规则是对博弈作出具体规定的集合。它包含对参与者行动顺序的规定、当某个参与者行动时他所知道的信息、有什么样的行动可供选择、选择之后会得到什么样的结果，等等。

(3) **结果(outcome)**。结果是所有参与者的每一个可能的行动组合最后会出现的情形。

(4) **收益(payoff)**。收益是在可能的每一个结果上，参与者的所得和所失。用经济学的话来说，就是在所有可能的结果上参与者的偏好是什么。这意味着，博弈的每一位参与者会在不同的结果之间进行比较，以寻求最好的收益。

一般而言，有什么样的结果就会有什么样的收益，因而在分析博弈的时候可以把结果和收益看作是一回事(这在分析上有好处)。不过严格来说，收益是结果

① 实际上自然是可以省略的，只是为了便于说明而保留。

的一个函数(payoff = f(outcome))。例如囚徒困境,如果囚徒1招认,囚徒2沉默,结果是囚徒1释放,囚徒2入狱9个月;囚徒1的收益为0,囚徒2的收益为 -9。

【例1.6】 石头、剪刀、布

(1) 参与者:孙悟空和猪八戒。通常用阿拉伯数字来代表,孙悟空定义为1,猪八戒定义为2。

(2) 博弈规则:双方同时出招,不得耍赖;石头赢剪刀、剪刀赢布、布赢石头。

(3) 结果:两种可能——(孙悟空休息,猪八戒找水)或是(孙悟空找水,猪八戒休息)。

(4) 收益:二人休息得到正的效用,比如为 $+U$;找水得到负的效用,比如为 $-U$。所以二人均偏好休息甚于找水,即休息总是比找水要好。

【例1.7】 诺曼底登陆

(1) 参与者:盟军和德军,分别定义为1和2。

(2) 博弈规则:盟军和德军彼此不知道对方的计划,如果两军相遇,盟军惨败;如果不相遇,盟军胜利。

(3) 结果:两种可能——(盟军胜利,德军惨败)或是(盟军惨败,德军胜利)。

(4) 收益:盟军胜利得到的收益为 $+\infty$,德军惨败得到的收益为 $-\infty$;盟军惨败得到的收益为 $-\infty$,德军胜利得到的收益为 $+\infty$。因而盟军和德军都希望胜利而不是惨败。

【例1.8】 烛光晚餐

假设有一对情侣,各自在不同的地点工作。下班后,双方由于客观原因无法联系上,但是他们希望能够在一起共进晚餐。他们常去吃饭的地点有两处,一是利顺德,一是喜来登。如果他们能碰在一起,双方共进烛光晚餐,感到非常幸福;如果遇不到一起,将各自在闷闷不乐中进餐。

(1) 参与者:情侣1和情侣2。

(2) 博弈规则:两个参与者分开,并且不能通信。他们必须各自作出在何处就餐的决定。

(3) 结果:要么共进晚餐,要么黯然神伤。

(4) 收益:共进晚餐得到的效用为 $+U$,分开进餐得到的效用为 $-U$。因而情侣1和情侣2偏好共进晚餐胜于黯然神伤。

任何一个博弈都具有这四个要素。对这四个要素进行科学的描述和表达,就为博弈的表达式。在博弈论中,博弈的表达式只有两种——基本式和扩展式。下面分别加以介绍。

㊥ 习题1.9 你能根据中国电信市场的现状,将中国移动和中国联通之间的竞争概括为一个博弈模型吗?它的四个要素是什么?

二、博弈的基本式(normal form)

如果用参与者、策略和偏好(即收益函数)来表述一个博弈,就称为博弈表达的基本式,也称为博弈的策略式(strategic form)。下面用囚徒困境来说明如何把一个博弈转变为基本式。根据前面的已知条件,写下囚徒困境的四个要素。

(1) 参与者:囚徒1和囚徒2。

(2) 博弈规则:坦白从宽,抗拒从严。

(3) 结果:要么1人释放,1人人狱9个月;要么两者都人狱6个月;要么两者人狱1个月。

(4) 收益:释放收益为0,人狱9个月收益为-9,人狱6个月收益为-6,人狱1个月收益为-1。

转变为基本式:

(1) 参与者集合:定义为$I=\{1,2\}$。1表示囚徒1,2表示囚徒2。

(2) 策略集合:定义为$S_i=\{$招供,沉默$\}$,$i=1,2$。S_i又称为策略空间。策略$s_{11}=$招供,$s_{12}=$沉默,$s_{21}=$招供,$s_{22}=$沉默。第一个下标表示第几个参与者,第二个下标表示第几个策略,例如$s_{11}=$招供,表示囚徒1的第一个策略为招供。其他以此类推。

(3) 偏好:对于囚徒1而言,从最好到最差依次排序的话,最好的是(招供,沉默)(即自己招供,对方沉默),其次是(沉默,沉默)(两个都沉默),然后为(招供,招供)(两个都招供),最差为(沉默,招供)(即自己沉默,对方招供)。对于囚徒2,偏好的排序为(沉默,招供),(沉默,沉默),(招供,招供),(招供,沉默)。如果用收益函数来表示,那么给定所有参与者的策略组合,就得到相应的收益。策略组合就是由每一个参与者的策略组成的向量。收益函数使得策略组合和收益之间建立了一种函数关系,定义为$u_i(s_{1j},s_{2j})$,$j=1,2$表示第几个策略。例如,$u_1(s_{11},s_{21})=-6$,即囚徒1招供,囚徒2招供,囚徒1得到的收益为-6(6个月监禁)。考察所有的策略组合,通过收益函数就能得到不同的收益,将它们列成矩阵形式,就得到图1-4。因而在博弈论中,又把博弈的矩阵表示等同于基本式。

(易) 习题1.10　请你将烛光晚餐博弈表示成博弈的基本式。

为什么把包含上述三个要素的表达式称为“基本”式?它“基本”在什么地方?孤立地解释显然不好理解,而结合博弈的扩展式表达就非常容易说明。

三、博弈的扩展式(extensive form)

表达博弈的第二种方法就是博弈的扩展式。如果详细地描绘出一个博弈的参与者、参与者的策略、参与者的行动顺序及其行动时所拥有的信息、可能的结果

和收益等细节，就称为博弈的扩展式。扩展式之所以称为“扩展”就是因为它比基本式“详细”。基本式中许多省略的细节（这就是为什么要把它称为基本式的根本原因），在扩展式中都得到了详细的说明。还是以囚徒困境为例，前面博弈矩阵（基本式）表示的囚徒困境也可以用博弈树（扩展式）来等价地表示，如图 1－7 所示。

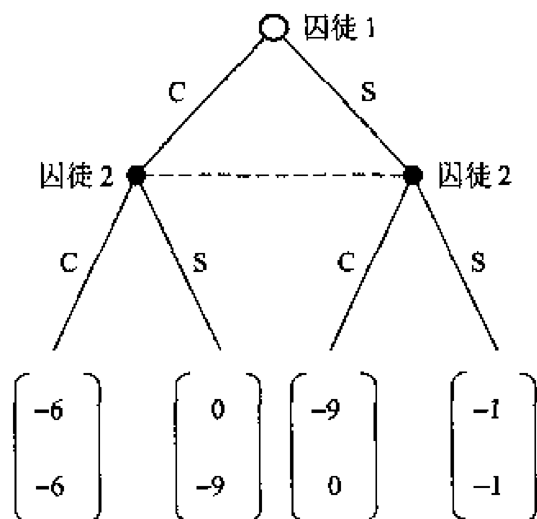


图 1－7　囚徒困境的扩展式

这里不妨假设囚徒 1 首先行动，囚徒 2 随后采取行动（也可以假设囚徒 2 首先行动，然后囚徒 1 采取行动，这表明基本式转化为扩展式的形式不是唯一的）。因面在第一个节点标注囚徒 1，表示囚徒 1 第一个行动。第一个节点称为起始点，通常用空心点表示。它有两个分支，分别代表可采取的两种行动——招供（用 C 表示）和沉默（用 S 表示）。在第二节点和第三节点标注囚徒 2，表示囚徒 1 行动后，轮到囚徒 2 行动。囚徒 2 也有两种可采取的行动——C 和 S。之后博弈结束，囚徒 1 和囚徒 2 分别得到各自的收益，第一个数值表示囚徒 1 的收益，第二个数值表示囚徒 2 的收益。这里囚徒 2 所处的两个节点被用虚线连接了起来，表示囚徒 2 在行动的时候，并不清楚囚徒 1 采取的行动是 C 还是 S，当然囚徒 1 也不清楚囚徒 2 采取什么行动（因为他在囚徒 2 前采取行动），因而就好比他们俩同时行动（正是这一点使得它与基本式等价）。打个比方，囚徒 1 首先被警方抓获，随后囚徒 2 也被警方抓获，囚徒 1 和囚徒 2 不存在串谋的可能。显然囚徒 2 不知道囚徒 1 是否招认，囚徒 1 同样也不知道囚徒 2 是否招认。因而这个博弈既可以用图 1－4 的博弈矩阵——基本式来表示，也可以用图 1－7 的博弈树——扩展式来表示。

显然，博弈树要比博弈矩阵详细得多。它清楚地表明了参与者的行动顺序（虽然这里不重要）、拥有的信息、信息的特征[①]、策略的本质等。例如，在基本

① 随着学习的深入，可以看到所有的完全信息静态博弈都是信息不完美的博弈。

式中,囚徒1和囚徒2的策略和行动是同一的,都为招认和沉默。如果将图1-7中连接囚徒2所处的节点的虚线取消(表示囚徒2知道囚徒1的行动),那么在扩展式中,囚徒1的策略为招认和沉默,行动和策略是同一的,囚徒2的策略则变为4个,即(如果囚徒1招认,那么招认),(如果囚徒1招认,那么沉默),(如果囚徒1沉默,那么招认),(如果囚徒1沉默,那么沉默),行动和策略完全不同。扩展式清楚地揭示出了策略的本质,即策略是随情况而定的完整的行动计划。行动并不等同于策略,策略是行动的指南,而行动只不过是策略的实践。在基本式中,囚徒2的策略就是行动,为招认和沉默,并没有体现随囚徒1策略的变化而变化这样的特征。也许有人会产生这样的疑问,是不是博弈的扩展式要比基本式"好"?从理论上来讲,任何基本式都可以转化为扩展式,而任何扩展式也可以转化为基本式①。用哪一个表达式关键是看哪一个方便。通常用基本式来表示静态博弈较为方便,而用扩展式来表示动态博弈则较为方便。

【例1.9】 国共合作

抗日战争时期,为了抗击侵略,国共进行了第二次合作,但双方时有冲突。针对这种情况,中国共产党提出了"人不犯我,我不犯人;人若犯我,我必犯人"的自卫原则。可以将国共合作概括成一个博弈,并用扩展式表达出来。我们会看到"人不犯我,我不犯人;人若犯我,我必犯人"的自卫原则是最优的策略,它既维护了国共合作的大局,又防止了国民党的得寸进尺。首先,博弈参与者为国民党和共产党。其次,国民党首先行动,有两种行动可以选择:一是不合作,二是合作。共产党在观察到国民党的行动后,选择相应的行动,也是两个:一是不合作,二是合作。最后,只要有一方选择不合作,那么国共合作就会破裂。有三种情况会使得国共合作破裂:一是国民党挑衅,而共产党合作,那么就有可能导致国民党的挑衅变本加厉,国共合作难以为继;二是国民党合作,而共产党不合作;三是国民党和共产党都选择不合作。三种情况都会导致国共合作破裂,而这是侵略者最希望看到的。由于选择合作的一方既受到日本侵略者的进攻,又受到不合作方的挑衅,因而损失要大,收益为-9,而挑衅方的收益为-5,如果国共双方都选择不合作,那么收益都为-1。上面的收益都为负,原因在于国共合作破裂对侵略者最有利,对国家民族最不幸。如果双方都选择合作,那么抗战就有了希望,对双方来说都是好事,收益为+1。具体的情况如图1-8所示。

① 对于基本式和扩展式是否等价,理论界仍然存在争论。但从直觉来说,越复杂的扩展式,通过越"详细"地定义策略,就可以使得基本式无限逼近扩展式,因而基本式和扩展式是等价的。不过,用哪一个表达式的标准主要是看它能否达到要求和是否简便。

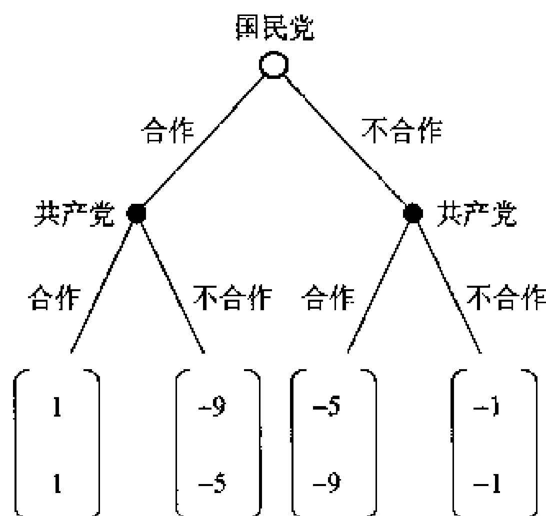


图 1-8　国共合作的博弈

从这个扩展式中可以看出，国共合作是最好的结局。如果国民党挑衅，那么共产党最好的做法就是进行反击（不合作）。在这个博弈中，共产党存在四种可能的选择，按偏好排序如下：（如果国民党合作，那么合作），（如果国民党不合作，那么不合作），（如果国民党合作，那么不合作），（如果国民党不合作，那么合作）。其中的（如果国民党合作，那么合作）和（如果国民党不合作，那么不合作）这个策略比另外三个策略要优，翻译成文学语言就是"人不犯我，我不犯人"和"人若犯我，我必犯人"的自卫原则。对于博弈的基本式和扩展式，在后面的章节中还会进行正式而严格的讨论。

四、信息和顺序

在一个博弈中，信息是至关重要的，孙子兵法中的"知己知彼，百战不殆"说的就是这个意思。掌握的信息不一样得到的结果也必然不一样，因为信息是作出正确决策的关键。有些信息对于博弈参与者而言是公共信息，而有些信息则只有自己知道。如果所有重要的信息都是**公共信息**（common knowledge），那么这样的博弈就称为完全信息博弈。如果一些重要的信息不是公共的，只有当事人知道，或者说不是所有的博弈者都知道，这样的博弈就称为不完全信息博弈。在博弈论中，对博弈的完全信息和非完全信息有着严格的定义：如果所有参与者在给定任意策略组合下，每一个参与者的收益都是确定的，那么就是**完全信息博弈**。如果至少有一个参与者的收益是不确定的（不确定是指参与者主观认为收益具有多种可能性）①，那么就是**非完全信息博**

① 如果把不确定性处理成服从一定的概率分布，实际上就等于把非完全信息处理成了完全信息。从这一点来看，完全信息博弈下的混合策略就具有了一点点不确定性。在后面的章节中会专门进行讨论。同时，随着学习的深入，读者会逐步明白这一点。

弈。上述这个定义等价于每个参与者的收益函数是否为公共信息，如果每个参与者的收益函数为公共信息，显然每一个参与者的收益必然是确定的。收益不确定是由于有些参与者只能“猜”他人或自己可能的收益。为什么“猜”？因为参与者的收益函数不是公共信息。公共信息也译为**公共知识**，它定义为：你知道这件事，他知道这件事，你知道他知道这件事，他知道你知道这件事，你知道他知道你知道这件事，你也知道他知道你知道他知道这件事……以至无穷。通俗地说，就是博弈的任意两个参与者双方彼此心里都清楚地知晓对方知道此事。与公共信息相对应的就是私人信息。私人信息就是一部分人知道，或是所有人不知道但自然知道。非完全信息博弈比完全信息博弈复杂，分析起来也较困难，但非完全信息较符合实际情况，因而适用范围更广，理论意义更大。

博弈参与者行动的顺序也很关键。他们可以同时行动，也可以有先有后。如果博弈双方同时行动（在博弈论中，所谓同时行动是指一方在做出行动时并不清楚对手是否已经做出了行动），称之为**静态博弈**。如果一方在做出行动时，知道对手已经做出了行动（可能不知道具体行动是什么，就像图 1－7 囚徒困境），称之为**动态博弈**。例如，石头、剪刀、布游戏就是一个静态博弈，而下象棋则是一个动态博弈。从这里可以看出所谓动态与静态的划分与时间先后无关，而与博弈的规则有关。从上面这两点出发，可以把博弈分为四类：完全信息静态博弈、完全信息动态博弈、非完全信息静态博弈和非完全信息动态博弈。相应的，就有四个均衡概念（见表 1－1）。

表 1－1　博弈的分类和均衡概念

	完全信息	非完全信息
静态	完全信息静态博弈 （纳什均衡）	非完全信息静态博弈 （贝叶斯均衡）
动态	完全信息动态博弈 （子博弈精炼均衡）	非完全信息动态博弈 （序列均衡）

四个博弈之间的关系还可用图 1－9 来描绘。

从图 1－9 还可以看到，一般均衡只不过是非完全信息动态博弈的一个点而已。如果把一般均衡比喻为鲁宾孙荒岛，那么非完全信息动态博弈对应的则是整个世界，博弈论就像一艘航船，使得鲁宾孙可以离开荒岛探索世界。

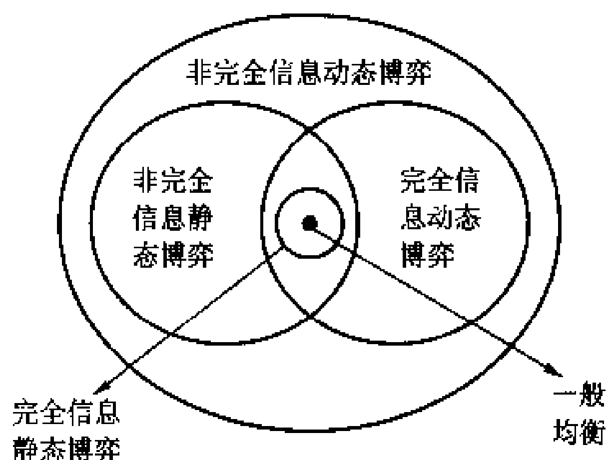


图 1-9　四个博弈模型关系图

公共知识信息（公共知识）的进一步讨论

假设有两个参与者，那么根据公共知识的定义，我们有：(a)甲知道这件事；(b)乙知道这件事；(c)甲知道(b)；(d)乙知道(c)……以至无穷。

这个例子充分反映出公共知识的特点——它可以不断迭代定义以至无穷。虽然公共知识难以定义，但却不影响我们使用它，就像数学中的1这个数一样，虽然难定义却不影响使用。实际上，在博弈中所有公共信息都是事先给定的（什么时候什么人知道什么样的信息都是事先规定好的），也就是说博弈的结构是外在于模型的。如何来理解这种公共知识和一般信息的关系呢？假设一个老师讲授博弈论，上课时间、地点、讲授学时、考试时间和地点这些就是所谓的公共知识，因为它们被公布在学校的网站上，并被所有学生看到，到考试的时候所有学生都会来，没有漏掉一个（这表明公共知识成立，否则就会有学生不来）。由于每个学生在学习博弈论的过程中掌握的信息不同，就会导致不一样的成绩（即博弈中的收益）。所以说，公共知识不会改变收益，但私人信息却会。

在博弈论中，有两个最基本的公共知识（理论假设）：理性的公共知识和先验概率的公共知识。理性的公共知识是说博弈的参与者都知道彼此是理性的，都追求自身利益的最大化。先验概率的公共知识表明自然(nature)对于所有人都是公平的，而不会偏向任何人。如果这两个假设不成立，那么将无法对博弈进行分析。

关键词

博弈　博弈论　策略的相互依存性　策略函数　模型　偏好理性　收益函数　单调变换　理性选择原理　决策主体　行为　制度结构　均衡　市场失灵　逆向选择　道德风险　搭便车行为　隐集团　极小极大解　博弈的四个要素　博弈的基本式　博弈的扩展式　博弈树　公共信息　私人信息　完全信息和非完全信息博弈　同时行动和静态博弈　动态博弈

内容要点

1. 博弈的实质是存在利益冲突的参与者总是试图“智胜对手”以实现自身利益的最大化，这使得博弈双方（或多方）的策略具有相互依存性。而博弈论就是针对这种策略依存性来探索各种博弈模型中的可能出现的均衡，并对各种均衡进行规律性的总结和概括，从而帮助人们理解各种社会现象。

2. 博弈论又可看作是由一系列模型所组成，博弈论理解各种冲突现象的方法就是将其抽象成一个模型（当然不能乱抽象）。这个模型通常包含着三个部分：理论前提、逻辑推理和理论结论。在发展中，博弈论深受经济学思想的影响，理性选择理论构成了博弈论最为核心的立足点，均衡则成为博弈论研究的中心。反过来，博弈论的发展也使得经济学，乃至其他社会科学都发生了一场革命，使得社会科学真正名至实归，因为如果离开了人的社会性，任何社会研究都是苍白的。

3. 由于存在垄断、公共产品、外在性、逆向选择和道德风险，使得一般均衡难以成立，完美的市场机制只不过是一种理想状态。没有最好，那能不能找到更好的？随着博弈论的发展我们有了回答这个问题的能力。从目前来看，博弈论对于理解市场失灵，以及找到解决之策具有难以替代的理论意义。将来可能会产生比博弈论更好的理论，但“在长期我们都死了”[①]。

4. 构建一个博弈模型离不开必要的材料，这些材料有四：参与者、规则、结果和收益。博弈的表达式又分为两个：基本式（策略式）和扩展式，基本式通常可以用博弈矩阵来直观表示，而扩展式则用博弈树表示。

5. 博弈论研究的难点主要体现在两点上：一是信息，二是顺序。根据信息

① 这是凯恩斯针对当时经济学界认为在长期内市场会自动恢复均衡，因而不需要政府干预所说的话。这里只是借用凯恩斯的话。

的完全与否,博弈可以分为完全信息博弈和非完全信息博弈。信息完全与否,则取决于每一个博弈参与者在任意策略组合下对其收益是否确定。而判断顺序的根据则是博弈参与者是否同时行动。同时行动与时间的先后无关,而与博弈的信息和规则有关。根据博弈参与者行动的顺序,可将博弈分为静态博弈和动态博弈。因而根据信息与顺序的组合关系,得到了四个博弈模型,即完全信息静态博弈、完全信息动态博弈、非完全信息静态博弈和非完全信息动态博弈,相应的有了四个均衡概念。其中非完全信息动态模型是最复杂的模型,对应真实的客观世界,其他模型都可以看作是它的特例。

进一步阅读

关于博弈论的发展史及其介绍可参考《帕尔格雷夫经济学大辞典》的博弈论词条(奥曼,1996);适合初学者的一般性介绍可参考张维迎(1996,第0章);有关理性选择理论则可以参考 Kreps(1990);有关理性选择理论局限性的讨论则见 Rabin(1998)。

第二章

完全信息静态博弈

什么是完全信息？是不是所有当事人都无所不知，还是所有信息都是公共信息？显然要作出以上这些判断是非常困难的，因为不同的情景要求的信息是不一样的。在博弈论中对完全信息的定义非常简单，但很严格，即如果所有参与者对他们的收益的信念都是确定的（不存在不确定性），那么这个博弈就是完全信息的。也就是说，所有参与者的收益函数是确定的是公共信息。只要有一个参与者的收益函数不是确定的，那么这个博弈就属于非完全信息。例如在石头、剪刀、布这个博弈中，由于孙悟空和猪八戒的收益函数是公共信息，所以当猪八戒出策略 x（例如剪刀），孙悟空出策略 y（例如布）时，猪八戒和孙悟空就能计算出他们各自将会得到多少的收益（例如 $+U$ 和 $-U$），即收益是确定的。因而，完全信息不仅意味着收益函数是确定的，而且还是公共信息。

静态博弈是指博弈的参与者都"同时行动"。所谓同时行动是指博弈中的任何一方在采取行动时，并不知道对手采取了什么样的行动，它与参与者行动的时间先后无关，而与博弈的顺序有关。如果一个博弈同时具备完全信息和静态的特征，即为完全信息静态博弈。

完全信息静态博弈是博弈论中最简单的博弈模型，但却是博弈论产生的基础。纳什（Nash）在1950年发表的不足千字的论文[①]中，首次在思想上证明了完全信息静态博弈存在着一般的均衡解，这个一般均衡解不过是博弈数学模型中的不动点。正是有了纳什均衡这个概念，博弈论的研究才真正走上了正确的方向。为了纪念纳什的贡献，后人把这个均衡概念称为纳什均衡。

在这一章中，主要介绍博弈表达的基本式，并使读者逐步掌握如何将一个博弈转换为基本式，并从简到难地给出纳什均衡这一重要概念，最终目的是使读者能够运用所学的知识建立最简单的博弈模型并求解纳什均衡，同时也为后面章

① Nash J. "Equilibrium Points in n - Person Games". *Proceeding of the National Academy of Sciences*. 1956，36：48 - 49

节的学习打下基础。

第一节 博弈的基本式

完全信息静态博弈是最简单的一种游戏①形式。在这种游戏里，每个参与者同时出招一次，随即游戏结束。换言之，每一个参与者各有一组可用的策略（行动），他们选择某一行动后，游戏结束，而所有参与者都得到各自相应的收益。在这个游戏中，关键之处在于每一个参与者采取什么行动取决于对手采取什么样的行动，并形成一种相互的影响。这种形式的游戏可用三个基本要素来完整地加以描述。

（1）参加博弈的有哪些"人"。人之所以打引号，是因为有些参与者可能是企业、利益集团（例如卡特尔）、国家或是军事同盟。通常把这些参与者定义为阿拉伯数字 1,2,3,…,n，表示第一位参与者、第二位参与者等，一直到第 n 位参与者。所有参与者的集合定义为大写 N，即 $N=\{1,2,\cdots,n\}$。

（2）每一个参与者的可用策略。每一个参与者所有可用的策略集合定义为策略空间，用大写的 S_i来表示。例如，第 i 位参与者所有可用的策略为 x,y,z，那么 $S_i=\{x,y,z\}$。具体的策略定义为 s_{ij}，它表示第 i 位参与者的第 j 个策略。显然 $s_{i1}=x,s_{i2}=y,s_{i3}=z$。策略空间是参与者所可以选择的策略，不在策略空间之内的策略是参与者无法选择的策略。如果把所有参与者选择的策略按顺序排成一行或列（在数学上称为向量），就为策略组合。例如，一个博弈有三位参与者，某个策略组合就为 $s=(s_{1j},s_{2k},s_{3h})$，即第 1 位参与者所用的策略为他的策略空间中的第 j 个策略，第 2 位参与者所用的策略为他的策略空间中的第 k 个策略，第 3 位参与者所用的策略为他的策略空间中的第 h 个策略。从数学来看，策略组合只不过是策略集合 $S_1\times S_2\times S_3$ 中的一个元素，即 $s\in S_1\times S_2\times S_3$。它表示第 1 位参与者、第 2 位参与者和第 3 位参与者分别在各自的策略空间中选择某一策略，然后形成一个三维向量 $s=(s_{1j},s_{2k},s_{3h})$。如果博弈有 n 位参与者，那么策略组合 $s\in S_1\times S_2\times\cdots\times S_n$，所有参与者的策略空间集合就用 $S=S_1\times S_2\times\cdots\times S_n$ 来表示。在数学中，$S=S_1\times S_2\times\cdots\times S_n$ 又称为 n 重笛卡尔积。有时也将某个策略组合写作 $s=(s_i,s_{-i})$，其中 s_i 为第 i 个参与者的策略，s_{-i}则表示除 i 以外其他参与者的策略组合，即 $s_{-i}=(s_1,\cdots,s_{i-1},s_{i+1},\cdots,s_n)$。

① 为了轻松一些，不至于太严肃，有时会用游戏来替代博弈一词，在英语中"博弈"（game）一词的意义就是游戏。

(3) 偏好和收益函数。偏好就是每一个参与者根据好坏对所有结果进行的排序。由于收益函数完全等价于偏好,而且更为简单明了,因而通常用收益函数来表示偏好。但这也有缺点,缺点就是会用到数学,对于"恐数症"的读者而言可能有所不便,不过一旦适应就会发现它有很多优点。开始的时候,我们会既用偏好也用收益函数,然后慢慢过渡到只用收益函数(如有必要也会用偏好),到时不再另加说明。每一个参与者得到多少收益取决于所有参与者的策略组合,而收益函数表示的是在什么样的策略组合下每位参与者得到什么样的收益,因而策略组合与收益之间的关系,就可以用收益函数$u(s_1,s_2,\cdots,s_n)$来表示。

有了上面这三个要素,就可以定义博弈的基本式。

定义 2.1 博弈表达的**基本式**(**或策略式**)由参与者集合 N、策略空间 S 和收益函数 u 三个要素组成,即 $G=\{N,S,u\}$,其中 $N=\{1,2,\cdots,n\}$,$S=\{S_1,S_2,\cdots,S_n\}$,$u=\{u_1,u_2,\cdots,u_n\}$。收益函数 $u_i:S\rightarrow R$,它表示第 i 位参与者在不同策略组合下所得到的收益。有时也把博弈表达的基本式写作 $G=\{S_1,S_2,\cdots,S_n;u_1,u_2,\cdots,u_n\}$或简写为$\{S,u\}$[①]。

(易) 习题 2.1 博弈的基本式包含哪些要素?为什么又称为策略式?

实际上,通俗的理解就是参与者、策略和偏好(收益函数)构成了博弈的基本式。

下面用一些例子来说明如何将一个博弈转化为博弈的科学表达式——基本式。

【例 2.1】 石头、剪刀、布

(1) 参与者集合:1 表示孙悟空,2 表示猪八戒,$N=\{1,2\}$。

(2) 策略空间:孙悟空的策略空间为 $S_1=\{$石头,剪刀,布$\}$,猪八戒的策略空间为 $S_2=\{$石头、剪刀、布$\}$。策略 $s_{11}=s_{21}=$石头,$s_{12}=s_{22}=$剪刀,$s_{13}=s_{23}=$布。

(3) 偏好和收益函数:孙悟空的偏好最好是(石头,剪刀)[②]、(剪刀,布)和(布,石头),其次为(石头,石头)、(剪刀,剪刀)和(布,布),最差为(石头,布)、(剪刀,石头)和(布,剪刀)。对于猪八戒的偏好,最好是(剪刀,石头)、(布,剪刀)和(石头、布),其次是(石头,石头)、(剪刀,剪刀)和(布,布),最差是(布,石头)、(石头、剪刀)和(剪刀,布)。等价地,定义 $u_1(s_{1j},s_{2k})$和 $u_2(s_{1j},s_{2k})$分别为孙悟空和猪八戒的收益函数。根据收益函数,可写出孙悟空和猪八戒在所有策

① 在完全信息静态博弈中,参与者的策略就是行动,如果定义 A 为行动空间集合,那么 $G=\{A,u\}$。

② 这里约定俗成地认为第一项表示第一个参与者的策略,第二项表示第二个参与者的策略,以此类推,不再说明。

略组合下可得的收益。例如，$u_1(s_{11},s_{21})=u_2(s_{11},s_{21})=0$，即孙悟空出石头，猪八戒也出石头，两者平局，收益都为0；$u_1(s_{11},s_{22})=+1$，$u_2(s_{11},s_{22})=-1$，即孙悟空出石头，猪八戒出剪刀，孙悟空休息，收益为+1，而猪八戒找水，收益为-1。当考察了所有的策略组合，并得到相应的收益后，可将其列成一个矩阵，如图2-1。由于博弈矩阵清楚地概括出了博弈的以上三个基本要素，因而博弈的基本式通常用博弈矩阵来表示。

		猪八戒		
		石头	剪刀	布
	石头	0,0	+1,-1	-1,+1
孙悟空	剪刀	-1,+1	0,0	+1,-1
	布	+1,-1	-1,+1	0,0

图2-1 石头、剪刀、布博弈矩阵

不过，如果参与者超过三人，策略数增多，那么用博弈矩阵来表示博弈的基本式显然就不方便了。通常，一个博弈的策略组合的数量等于k^n，k为每一个参与者拥有的策略数量，而n为参与者数量①。显然，随着策略数增多或是参与者增多，博弈可能出现的策略组合呈几何数增长。因而从这一点来看，并不是所有的博弈基本式都可用博弈矩阵来概括。

【例2.2】 囚徒困境

(1) 参与者集合：囚徒1定义为1，囚徒2定义为2，$N=\{1,2\}$。

(2) 策略空间：囚徒1的策略空间$S_1=\{$沉默，招供$\}$，囚徒2的策略空间$S_2=\{$沉默，招供$\}$。策略$s_{11}=s_{21}=$沉默，$s_{12}=s_{22}=$招供。

(3) 偏好和收益函数：囚徒1的偏好，从好到差排序为(招供，沉默)(即对方沉默他招供)，(沉默，沉默)，(招供，招供)，(沉默，招供)。囚徒2的偏好与此类似，不再赘述，只是注意不要把顺序搞错，第一项为囚徒1的策略，第二项为囚徒2的策略。定义$u_1(s_{1j},s_{2k})$和$u_2(s_{1j},s_{2k})$分别为囚徒1和囚徒2的收益函数。所有的收益如下：

$$u_1(s_{11},s_{21})=-1 \quad u_1(s_{11},s_{22})=-9 \quad u_1(s_{12},s_{21})=0 \quad u_1(s_{12},s_{22})=-6$$
$$u_2(s_{11},s_{21})=-1 \quad u_2(s_{11},s_{22})=0 \quad u_2(s_{12},s_{21})=-9 \quad u_2(s_{12},s_{22})=-6$$

显然，

$$u_1(s_{12},s_{21})=0>u_1(s_{11},s_{21})=-1>u_1(s_{12},s_{22})=-6>u_1(s_{11},s_{22})=-9$$
$$u_2(s_{11},s_{22})=0>u_2(s_{11},s_{21})=-1>u_2(s_{12},s_{22})=-6>u_2(s_{12},s_{21})=-9$$

① 如果参与者拥有的策略数不相同，那么就是$k_1\times k_2\times\cdots\times k_n$，下标表示第几位参与者。

与囚徒1和囚徒2的偏好完全一致。

将所有收益列成矩阵形式，得图2-2。

		囚徒2 沉默	囚徒2 招供
囚徒1	沉默	-1, -1	-9, 0
囚徒1	招供	0, -9	-6, -6

图2-2 囚徒困境博弈矩阵

㊣ 习题2.2 人民公社。在联产承包责任制实行之前，中国农村实行的制度为人民公社，即大伙一起劳作，一起到公社食堂吃饭，人们可以选择努力劳作也可以偷懒（是努力还是偷懒很难判断，全凭自觉）。由于干好干差都一样，结果人民公社很快就破了产（指人民公社被吃空，而不是指法理上的破产）。这里我们试图建立一个博弈模型来说明人民公社当时的情景。你能建立这个模型吗？它的基本式是什么？它与囚徒困境存在什么样的联系？（提示：假设有两个参与者（即公社社员，实际上 n 个参与者也一样），有两种策略：努力和偷懒。）

【例2.3】 斗鸡博弈

在美国，一些飞车党党徒为了表示勇敢，通常由两个人分别驾驶两辆摩托车（或是汽车）从相向的方向急速对撞。如果有一方怕死而首先让道，那么让道的人就被称为小鸡（chicken）。如果两人都同时让道，两人都被称为小鸡。被称为小鸡的人，在飞车党内地位低下，备受歧视。而不怕死的人，会在飞车党内享受崇高荣誉，受到众人推崇。显然，在斗鸡博弈中，如果两人都不让道，有可能不死既残，收益为 $-\omega$；如果主动让道，又会受到大家的嘲笑，做人抬不起头，收益为 -10；如果不让道成为勇敢者，会受到大家尊敬，收益为 $+10$。现在把斗鸡博弈转化为博弈的基本式。

(1) 参与者集合：党徒1定义为1，党徒2定义为2，$N=\{1,2\}$。

(2) 策略空间：党徒1的策略空间 $S_1=\{$让，撞$\}$，党徒2的策略空间 $S_2=\{$让，撞$\}$。策略 $s_{11}=s_{21}=$让，$s_{12}=s_{22}=$撞。

(3) 偏好和收益函数：党徒1的偏好，从好到差排列为（撞，让），（让，让）和（让，撞），最差为（撞，撞）。党徒2的偏好类似，不再赘述。定义 $u_1(s_{1j},s_{2j})$ 和 $u_2(s_{1j},s_{2j})$ 为党徒1和党徒2的收益函数。所有的收益如下：

$u_1(s_{11},s_{21})=-10$ $u_1(s_{11},s_{22})=-10$ $u_1(s_{12},s_{21})=+10$ $u_1(s_{12},s_{22})=-\omega$

$u_2(s_{11},s_{21})=-10$ $u_2(s_{11},s_{22})=+10$ $u_2(s_{12},s_{21})=-10$ $u_2(s_{12},s_{22})=-\omega$

将所有收益列成矩阵形式，得图2-3。

		党徒2	
		让	撞
党徒1	让	-10, -10	-10, +10
	撞	+10, -10	$-\omega, -\omega$

图2-3 斗鸡博弈矩阵

【例2.4】 双头垄断

在一个寡头垄断的市场结构中只存在两个企业,它们生产完全同质的产品,两个企业的产品可以完全相互替代。在竞争中,两个企业选择价格,而让市场决定需求。每个企业在价格上有两种选择,要么选择高价,要么选择低价。如果两个企业都选择高价,那么它们能得利润1 000亿元;如果一个企业选择高价,而另一个企业选择低价,那么高价企业的利润为500亿元,而低价企业的利润为1 500亿元;如果两个企业都选择低价,那么利润都为700亿元。显然,两个垄断企业的策略(即价格)存在着相互依存性,彼此存在利益上的冲突,双方的利润都会受到对方所选价格的影响。垄断企业甲并不能随心所欲地操纵市场,而必须考虑到垄断企业乙所可能作出的反应。对于垄断企业乙也同样如此。现在将其转化为博弈的基本式。

(1) 参与者集合:垄断企业甲定义为1,垄断企业乙定义为2,$N=\{1,2\}$。

(2) 策略空间:垄断企业甲的策略空间为$S_1=\{$高价,低价$\}$,垄断企业乙的策略空间为$S_2=\{$高价,低价$\}$。策略$s_{11}=s_{21}=$高价,$s_{12}=s_{22}=$低价。

(3) 收益函数:垄断企业甲乙的收益函数$u(s_1,s_2)$实际上就是垄断企业的利润函数。收益即为利润。根据已知条件可知,垄断企业甲和垄断企业乙的利润函数分别为:

$u_1(s_{11},s_{21})=1\ 000$ $u_1(s_{11},s_{22})=500$ $u_1(s_{12},s_{21})=1\ 500$ $u_1(s_{12},s_{22})=700$

$u_2(s_{11},s_{21})=1\ 000$ $u_2(s_{11},s_{22})=1\ 500$ $u_2(s_{12},s_{21})=500$ $u_2(s_{12},s_{22})=700$

将所有收益列成矩阵形式,得图2-4。

		企业乙	
		高价	低价
企业甲	高价	1 000,1 000	500,1 500
	低价	1 500,500	700,700

图2-4 双头垄断博弈矩阵

㊁ 习题 2.3 这个模型实际上和囚徒困境本质上一样。如果理解成移动电话市场，那么我们会得到什么样的均衡解？显然均衡解和现实存在着矛盾的地方，即两个电话公司都维持着极高的价格。为什么会出现这种矛盾？（提示：要回答这一题需要用到重复博弈中的知识，在完全信息动态博弈中可以找到相关内容。）

【例 2.5】 猎鹿

在博弈论中有三个非常著名的例子，其中的囚徒困境、鹰派和鸽派我们已经接触过，而现在所接触的猎鹿模型就是第三个模型。猎鹿模型是卢梭在其著作《关于人间不平等的根源及其基础的讨论》(1755)中提到的一个故事。一群猎人共同猎鹿，如果有一个人分心，猎鹿就会失败，因而猎鹿成功依赖于每一个猎人的共同努力。现在我们将它变成一个博弈模型。假设存在 n 个猎人，他们共同猎鹿，在猎鹿的过程中，他们可以始终把注意力集中在猎鹿上，那么将最终能够捕获野鹿。如果有人分心去抓野兔，那么猎鹿将会失败。其基本式为：

（1）参与者集合：n 个猎人，$N=\{1,2,\cdots,n\}$。

（2）策略空间：每个猎人的策略空间为 $S_i=\{$鹿，兔$\}$，$\forall i\in N$①。

（3）偏好：如果所有其他猎人都猎鹿，那么第 i 个猎人最好的选择也是猎鹿，其收益不妨假设为 3；如果有任何一个人跑去猎兔，那么猎鹿肯定失败，还不如去猎兔，收益为 1；如果有人跑去猎兔，而第 i 个猎人仍然坚持猎鹿，那么猎鹿的人将两手空空，收益为 0。（这时用收益函数来表示偏好有点烦琐，所以不如直接用偏好简单。）将收益列成矩阵，见图 2－5。

		其他猎人	
		鹿	兔
猎人 i	鹿	3,3	0,1
	兔	1,0	1,1

图 2－5 猎鹿模型博弈矩阵

猎鹿模型虽然简单，但却很有启发性。例如，民主制度的建立和维护需要所有人的努力，光靠一两个人的努力显然难以成功，这个时候就会出现大家都去猎兔的情况（鲁迅先生的小说《药》就生动地描述了这种情况），如果再考虑到民主制度具有的外在性，民主制度的建立就更加困难。相较而言，破坏民主制度比破坏专制制度更加容易。例如，很多国家推行“民主”的结果，往往是贿赂选举、政

① ∀表示 any（任意一个的意思），它实际上是 A 的倒写。

治腐败、军阀混战、民不聊生。真正的民主需要的客观条件比专制多得多，并不像我们想象的那么简单和容易。

㊥ 习题2.4 和平与备战。有两个邻国，维持和平显然需要双方的共同努力，如果有一方积极备战，那么和平局面将难以维持。和平对双方最有利，如果双方都备战，那么耗力又耗物，对双方都不利。但是如果一方采取和平政策，而另一方积极备战，那么备战一方相对于和平一方将获得更多利益。请你建立该博弈模型，并给出其基本式。

有了对博弈的科学表述，随后的问题就是如何找出博弈的均衡解，而纳什均衡这个概念是博弈均衡解的基础。

第二节 纳什均衡

在介绍纳什均衡之前，我们先从简单的均衡概念入手，逐步深入，最终给出纳什均衡的定义。

一、严格优策略均衡与一般均衡

研究博弈理论，最重要的目的就是希望在面对一个博弈时，可以了解并预测博弈最后的结果。在学习经济学的时候，已经可以看到在均衡状态下，所有行为主体都没有积极性主动偏离均衡状态，即使偏离了均衡状态，各行为主体也会想方设法回到均衡状态，因为均衡对所有各方来说是最优的。这说明一个博弈最后的结果最有可能是一个均衡状态。在博弈中，最简单的均衡概念就是严格优策略均衡。

下面还是用囚徒困境来加以说明。在囚徒困境中，什么是博弈最可能出现的结果？或换句话说，博弈的均衡解是什么？通过图2－2，很容易找出，如果囚徒2选择沉默，那么囚徒1最优的策略是招供；如果囚徒2选择招供，那么囚徒1最优的策略还是招供。这说明无论囚徒2的策略是什么，囚徒1选择招供都要比选择沉默要优。对于囚徒2而言，同样的分析也成立。因而，囚徒困境最可能出现的结果是（招供，招供），没有哪个囚徒会主动偏离这个结果，因为偏离它就意味着要遭受更大的损失。例如，囚徒2偏离它，选择沉默，那么坐牢就不是6个月，而是9个月。

定义2.2 在博弈$G=\{S,u\}$中，在其他参与者任意[①]给定的策略组合下，即

① 在数学中，“任意”用$\forall$来表示，它为Any（任意）头一个字母的倒写，“存在一个”用$\exists$来表示，它为Exist（存在）头一个字母的倒写。

$\forall s_{-i}=(s_1,\cdots,s_{i-1},s_{i+1},\cdots,s_n)\in S_{-i}=(S_1,\cdots,S_{i-1},S_{i+1},\cdots,S_n)$，参与者 i 存在一个策略 s_i^* 使得对于 $\forall s_i\in S_i$，$s_i^*\neq s_i$，都有

$$u_i(s_1,\cdots,s_i^*,\cdots,s_n)>u_i(s_1,\cdots,s_i,\cdots,s_n)$$

就称 s_i^* 为参与者 i 的**严格优策略**。

在囚徒困境中，无论囚徒 2 的策略是什么，囚徒 1 选择招供策略所得的收益都要大于选择沉默策略，所以招供是囚徒 1 的严格优策略。对于囚徒 2 而言，无论囚徒 1 选择什么策略，囚徒 2 选择招供策略都要优于沉默策略，所以囚徒 2 的严格优策略为招供。显然（招供，招供）构成了囚徒困境的一个均衡解。通过对囚徒困境的分析，很容易得到一个一般性的结论，即如果一个博弈中的所有参与者都存在严格优策略，那么严格优策略组合一定是该博弈的唯一均衡解。

命题 2.1　如果一个策略博弈 $G=\{S,u\}$ 的每个参与者都存在严格优策略，那么严格优策略组合一定是该博弈唯一的均衡解。

证：假设一个博弈为 $G=\{S,u\}$，其中每一个参与者都存在严格优策略。根据定义可知严格优策略具有唯一性，如果不是就意味着参与者 i 至少有两个严格优策略，不妨设它们分别为 s_i^* 和 $s_i^\#$，根据严格优策略的定义，可知

$$u_i(s_i^*,s_{-i})>u_i(s_i^\#,s_{-i})$$

$$u_i(s_i^\#,s_{-i})>u_i(s_i^*,s_{-i})$$

显然这是自相矛盾的。因而严格优策略具有唯一性。由于严格优策略具有唯一性，因而每一个参与者假如不选择严格优策略，将意味着他得到的收益不为最优，显然这与参与者是理性的相矛盾。从而命题得证。

(易) 习题 2.5　命题 2.1 的必要条件。在一个两人博弈中，如果只有一个参与者存在严格优策略，那么是否存在均衡解？如果没有参与者存在严格优策略，那么是否存在唯一均衡解？

定义 2.3　在博弈 $G=\{S,u\}$ 中，如果对所有参与者 i 而言，都存在 s_i^* 为严格优策略（其中 $\forall i\in N$），那么策略组合 $s^*=(s_1^*,\cdots,s_i^*,\cdots,s_n^*)$ 就称为**严格优策略均衡**。

在有些情况下，优策略也可能形成均衡，而不一定必须是严格优策略，但却可能出现多重均衡。

定义 2.4　在博弈 $G=\{S,u\}$ 中，如果对参与者 i 而言，在其他参与者任意给定的策略组合下，即 $\forall s_{-i}=(s_1,\cdots,s_{i-1},s_{i+1},\cdots,s_n)\in S_{-i}=(S_1,\cdots,S_{i-1},S_{i+1},\cdots,S_n)$，存在一个策略 s_i^* 使得对于 $\forall s_i\in S_i$，都有

$$u_i(s_1,\cdots,s_i^*,\cdots,s_n) \geqslant u_i(s_1,\cdots,s_i,\cdots,s_n)$$

并且至少存在一个 s_i 使得

$$u_i(s_1,\cdots,s_i^*,\cdots,s_n) > u_i(s_1,\cdots,s_i,\cdots,s_n)$$

成立,就称 s_i^* 为参与者 i 的**优策略**。

如果参与者只存在优策略而不是严格优策略,那么一个可能的后果是博弈不存在唯一的均衡解,从而使得我们无法对博弈的结果作出明确的预测,极端的情况是博弈有可能出现无穷多的结果,这时理论的预测能力就大打折扣。

【例 2.6】 密封式二级价格拍卖

在一个拍卖市场上,有 n 个竞拍者,标的物由出价最高者得,但是出价最高者实际支付的不是其报价,而是出价第二高者的报价,因而被称为二级价格拍卖。不妨假设,标的物的价值对于竞拍者 i 而言为 v_i,其报价为 b_i, $i \in N$。如果获得标的物,那么净收益为 $v_i - sb_i$,其中 sb_i 为第二高价。该博弈的基本式为:

(1) 参与者集合:n 个竞拍者。

(2) 策略空间:每一个竞拍者的策略为其报价 $b_i \in [0, +\infty)$。

(3) 收益函数:对于每一个竞拍者而言,都希望 $v_i - sb_i$ 越大越好。

我们将证明每一个竞拍者都存在一个**优策略**,这个优策略就是真实报告自己的评价,即 $b_i = v_i$。

首先,考虑报价大于价值的情况,即 $b_i > v_i$。如果其他竞拍者的报价都小于等于 v_i,那么显然 $b_i > v_i$ 与 $b_i = v_i$ 的收益一样;如果其他竞拍者的报价小于 b_i 但大于 v_i,那么 $b_i > v_i$ 要比 $b_i = v_i$ 的结果差,因为 $v_i - sb_i < 0$;如果其他竞拍者的报价大于等于 b_i,那么 $b_i > v_i$ 与 $b_i = v_i$ 的收益一样。

其次,考虑报价小于价值的情况,即 $b_i < v_i$。如果其他竞拍者的报价都小于 b_i,那么显然 $b_i < v_i$ 与 $b_i = v_i$ 的收益一样;如果其他竞拍者的报价小于 v_i 但大于等于 b_i,那么显然 $b_i < v_i$ 要比 $b_i = v_i$ 的收益小(有可能其他竞拍者的报价刚好等于 b_i,这时竞拍者 i 有可能得到标的物,也有可能得不到,其收益的期望值要小于 $b_i = v_i$所获得的收益);如果其他竞拍者的报价大于等于 v_i,那么 $b_i < v_i$ 与 $b_i = v_i$ 的收益一样。

综上所述,真实报告自己的评价,即 $b_i = v_i$,就是二级价格拍卖中每一个竞拍者的优策略。

㊥ 习题 2.6 在上面的二级价格拍卖中,均衡是否唯一?如果不唯一,请证明之。如果价格不是密封的,并且依次出价,会出现什么情况?

在微观经济学中,已知证明一般均衡存在的一个不可缺少的重要前提是市场为完全竞争的市场结构。完全竞争的本质是所有的市场参与者——消费者和生产者都是价格的接受者,而不是价格的影响者。无论是消费者还是生产者,他的决策(商品的消费量和供给量)对市场价格的影响都趋于零,从而每一个参与者的策略都不会影响到其他参与者的策略,也就是参与者之间不存在策略依存性,而这种特征和严格优策略是完全相同的。对消费者而言,他是在给定商品价格、收入等条件下使自身的效用最大化,至于其他参与者如何行动对他的决策不存在影响;对生产者而言,他是在给定商品价格、要素价格、成本、技术等条件下使自身的利润最大化,至于其他参与者如何,同样与他不存在联系。因而完全竞争市场上的每一个参与者实际上选择的策略都是严格优策略,形成的一般均衡就是一个严格优策略均衡。如果有参与者选择的策略是优策略,而不是严格优策略,例如在完全竞争下厂商的利润为零时就会出现这种情况,那么一般均衡就可能是多重的。

一般均衡为严格优策略均衡,显然大大加深了我们对一般均衡和博弈论的理解。由于一般均衡只不过是博弈论中纳什均衡的一个特例,因而博弈论适用的领域就要比一般均衡所描绘的范围宽,这就是博弈论能够在所谓"市场失灵"——外在性、公共产品、垄断、道德风险、逆向选择的世界大展拳脚的原因,制度、产权这些传统经济学作为外生给定因素来处理的东西,反而成为博弈论研究中的考察对象。

在所有博弈中,存在严格优策略的博弈毕竟是少数,绝大多数博弈并不存在严格优策略(或优策略)。不过在有些博弈中,虽然没有严格优策略,但存在严格劣策略,通过剔除严格劣策略仍可能得到博弈的均衡解。

定义 2.5 在博弈 $G=\{S,u\}$ 中,s_i' 和 s_i'' 为参与者 i 的两个可行策略,即 $\exists s_i',s_i''\in S_i$。如果对于 $\forall s_{-i}\in S_{-i}$,有

$$u_i(s_i',s_{-i}) > u_i(s_i'',s_{-i}),$$

那么 s_i'' 就称为相对于策略 s_i' 的**严格劣策略**。

作为理性者,显然没有哪个参与者会选择严格劣策略,因而可以将严格劣策略从参与者的策略空间中剔除。通过不断地剔除严格劣策略就可能最终得到博弈的均衡解,这种方法就称为重复剔除严格劣策略法。下面用例子来加以说明。

【例 2.7】 上下一左中右

该博弈有两个参与者,参与者 1 的策略为上和下,参与者 2 的策略为左、中、右。不同策略组合下相应的收益见图 2-6。

		参与者2 左	中	右
参与者1	上	1,0	1,2	0,1
	下	0,3	0,1	2,0

图2-6 上下—左中右博弈矩阵

在这个博弈中,参与者1显然不存在严格优策略,也不存在严格劣策略。参与者2虽不存在严格优策略,但却存在严格劣策略。策略右相对于策略中就是一个严格劣策略。因为无论参与者1选择什么策略,参与者2选择中都要优于选择右。如果参与者是理性的,无论什么情况,参与者2绝对不会选择右,因而可以将右从参与者2的策略中剔除。从而得到图2-7。

		参与者2 左	中
参与者1	上	1,0	1,2
	下	0,3	0,1

图2-7 上下—左中博弈矩阵

从图2-7可以看到,对于参与者1而言,下是一个严格劣策略,可从参与者1的策略空间中剔除,最终得到图2-8。

		参与者2 左	中
参与者1	上	1,0	1,2

图2-8 上—左中博弈矩阵

显然(上,中)就是该博弈唯一的均衡解。

存在严格劣策略,自然也存在**劣策略**,它的定义类似于优策略的定义,请读者自己举一反三,这里不再说明。

㊥ 习题2.7 请你给出劣策略的定义。

重复剔除严格劣策略这种方法能不能一般地运用到劣策略呢?答案是否定的,因为重复剔除劣策略有可能将博弈的均衡解给剔除掉,只有在严格劣策略下,才不会出现这种情况。原因在于,严格劣策略无论在什么情况下都是严格不等式成立,而在劣策略下有些策略的收益是相等的。例如,在图2-9中,存在着

两个均衡解(A,A)和(B,B)[1]。

		参与者2 A	参与者2 B
参与者1	A	1,1	4,0
	B	0,4	4,4

图2-9 均衡解—劣策略

从直观上看,策略组合(B,B)要明显帕累托优于策略组合(A,A),但根据劣策略的定义,显然B对于任何一个参与者而言都是一个劣策略。如果运用重复剔除劣策略的话,就会将(B,B)这个均衡解给剔除掉,显然这是极不合理的。下面再看一个例子。

【例2.8】 投票

在某个乡村,奇数的村民投票选举村长,每人一票,获得多数选票的候选人获得村长职务。村长候选人有两位:A和B。有的村民喜欢A,而另一些村民喜欢B,自己喜欢的候选人获胜要好过自己不喜欢的候选人获胜。由于村民数为奇数,因而不会出现投票数一样的情况。其博弈的基本式如下:

(1) 参与者集合:奇数的村民。

(2) 策略空间:每一个村民其策略空间为{A,B}。

(3) 偏好。对于每个村民而言,自己喜欢的候选人获胜是最好的,不好的就是自己不喜欢的候选人获胜。

投票模型最具有启发性的地方就是,在有些均衡解中,一些参与者的策略是劣策略。根据上面的题意,很容易证明任一个村民选择投票给自己不喜欢的候选人是一个劣策略。因为给定其他村民的策略,如果该村民投票给自己不喜欢的候选人,那么只可能出现两种情况:一种是使得自己不喜欢的候选人获胜(他的投票就是使骆驼垮掉的最后一根稻草),显然结果要比投给自己喜欢的候选人要差;另一种就是不会改变选举结果,这时投票给不喜欢的候选人与投票给喜欢的候选人两者结果一样。所以,投票给自己不喜欢的候选人对于任意村民来说都是一个劣策略。根据均衡解的定义,所有村民投票给一个候选人显然是一个均衡解,这就证明了在该均衡解中,对于一些村民而言其策略是一个劣策略。

投票模型虽然简单,但却能够解释很多选举中出现的现象。例如,如果允许

① 这两个策略组合都是纳什均衡解,纳什均衡解是本节下一节部分内容的核心。由于说明劣策略的特性要用到纳什均衡的概念,所以有些教材又将这部分内容放在纳什均衡内容的后面。

村民弃权,那么投票率非常低就可能是一个均衡;由于每一个选民都认为自己的一票无关大局,因而就可能出现贿选的情况,即使这个候选人不为大多数人所喜欢也可能赢得选举。由于选举是民主制度的根基,选举制度中出现的缺陷就有可能使得民主名不副实,成为权贵的掌中玩物。

㉆ 习题 2.8　在投票模型中,考虑投票需要一定的成本(例如,到投票站投票需要一定的时间、金钱和精力),并允许选民弃权,请你写出该博弈的表达式,并找出一个均衡解(提示:参考纳什均衡的概念),并用其来解释为什么存在投票率过低和贿选的情况。

严格劣策略虽然要求的条件没有严格优策略那么苛刻(严格优策略要求对所有策略都要优,而严格劣策略只是相对于某个策略而言,这意味着存在严格优策略必然存在严格劣策略,存在严格劣策略却不一定存在严格优策略),但毕竟具有严格劣策略的博弈仍然是少数。对于更为一般的博弈,如何求出均衡解呢?这就引出了纳什均衡这个概念。

二、纳什均衡

在前面举了许多例子,例如诺曼底登陆、鸽派和鹰派、斗鸡博弈、猎鹿博弈等。在这些例子中,没有哪一个参与者拥有严格优策略或是严格劣策略,但从直觉上说,我们都认为这些博弈存在着均衡解。在这些博弈中,参与者的最优选择是什么呢?换句话说,均衡状态下的策略应具有什么样的特征?当每个参与者在决定行动时,一定会对其他参与者可能采取的行动形成一种信念(belief),正是这种信念决定了什么样的策略才是最优的,当所有的参与者的信念都达到一致时就出现了均衡。因而,要实现一个均衡,至少要满足:① 参与者的信念是正确的。如果信念有误,最优决策也就无从谈起。② 理性的参与者根据其信念来选择最优行动。沿着以上思路,我们就有:

在一个存在 n 个参与者的博弈中,如果策略组合 s^* 是一个均衡的话,那么给定 $n-1$ 个参与者的策略,某个参与者所选择的策略一定是最优的(否则就与理性人假设相矛盾);又由于参与者是任意选取的,因而上述特性对于每个参与者都成立。换言之,对于一个策略组合 s^*,如果保持其他参与者的策略不变,而任意一个参与者的策略都是最优的,那么策略组合 s^* 就是纳什均衡。

如果将以上内容用数学表示出来,就有如下定义。

定义 2.6　如果一个策略组合 $s^*=(s_1^*,\cdots,s_n^*)$ 是博弈 $G=\{S,u\}$ 的一个纳什均衡,那么对任一参与者 $i=1,2,\cdots,n$,对其任一个策略 $s_i\in S_i$,不等式

$$u_i(s_i^*,s_{-i})\geqslant u_i(s_i,s_{-i}) \tag{2-1}$$

成立。

记忆纳什均衡的一个诀窍就是将它归纳为一句话："最优对任一参与者的任一策略成立"。

显然，纳什均衡是判断一个策略组合是否为均衡的充分必要条件。利用它可以对各种各样的博弈进行分析，并找出均衡解。下面就利用纳什均衡来寻找前面提到的一些例子的均衡。

1. 鸽派和鹰派（见例 1.3）

该博弈存在两个纳什均衡：（鹰派政策，鸽派政策）和（鸽派政策，鹰派政策）。（鹰派政策，鸽派政策）之所以为均衡，在于如果给定美国的策略——鸽派政策，那么苏联的最优策略是鹰派政策；反过来，如果给定苏联的策略是鹰派政策，那么美国的最优策略是鸽派政策。该策略组合完全符合纳什均衡的定义，所以（鹰派政策，鸽派政策）就是该博弈的一个纳什均衡。对子策略组合（鸽派政策，鹰派政策）同样如此，给定美国的策略是鹰派政策，那么苏联的最优策略是鸽派政策；如果给定苏联的策略为鸽派政策，那么美国的最优策略为鹰派政策，所以（鸽派政策，鹰派政策）也是该博弈的一个纳什均衡。

2. 富人穷人博弈（见例 1.5）

该博弈存在唯一一个纳什均衡：（巡逻，不巡逻）。给定穷人的策略是不巡逻，那么富人的最优策略是巡逻；反之，如果给定富人的策略是巡逻，那么穷人的策略是不巡逻。所以，（巡逻，不巡逻）是富人穷人博弈的纳什均衡。

3. 斗鸡博弈（见例 2.3）

该博弈存在两个纳什均衡：（让，撞）和（撞，让）。给定党徒 2 不要命地撞过来，那么党徒 1 最优的策略是让，失去尊严总比命丧九泉要好；反之，如果党徒 1 贪生怕死，那么党徒 2 最优的策略就是撞。所以，（让，撞）就是斗鸡博弈的一个纳什均衡。另一个纳什均衡为（撞，让）。

4. 猎鹿（见例 2.5）

该博弈存在两个纳什均衡：（鹿，鹿）和（兔，兔）。给定其他猎人都抓鹿，那么猎人 i 最优的策略是抓鹿；反之，如果所有人都抓鹿，那么任一猎人最优的策略也只能是抓鹿。对子策略组合（兔，兔），分析同理。

5. 囚徒困境（见例 2.2）

囚徒困境存在唯一的纳什均衡：（招供，招供）。如果给定囚徒 2 选择招供，那么囚徒 1 的最优策略只能是招供；反之，给定囚徒 1 选择招供，囚徒 2 的最优策略也只能是招供。所以，（招供，招供）就是一个纳什均衡。

在囚徒困境中，（招供，招供）不仅是一个纳什均衡，而且还是一个严格优策

略均衡。那么,两者之间存在什么关系呢?实际上,严格优策略均衡必定是纳什均衡,但反之不成立。例如,鸽派和鹰派博弈、富人和穷人博弈、斗鸡博弈、猎鹿博弈都不存在严格优策略,但却存在纳什均衡。这表明纳什均衡概念包含着严格优策略均衡,它是一个比严格优策略均衡更一般的概念。

命题 2.2 如果博弈 $G=\{S,u\}$ 存在严格优策略均衡 $s^*=(s_1^*,\cdots,s_n^*)$,那么它一定是一个纳什均衡,但反之不成立。

证:利用反证法。

已知策略组合 $s^*=(s_1^*,\cdots,s_n^*)$ 是博弈 $G=\{S,u\}$ 的严格优策略,但不是纳什均衡。不妨假设纳什均衡为 $s^\#=(s_1^\#,\cdots,s_n^\#)$。根据严格优策略的定义可知,一定存在参与者 $i=1,2,\cdots,n$,有

$$u_i(s_i^*,s_{-i}^\#) > u_i(s_i^\#,s_{-i}^\#)$$

而根据纳什均衡的定义,一定有

$$u_i(s_i^\#,s_{-i}^\#) \geqslant u_i(s_i^*,s_{-i}^\#)$$

显然这是自相矛盾的,除非 $s^*=s^\#$。

要证第二个观点,不妨假设凡是纳什均衡一定是严格优策略均衡。只要找出一个反例,这个假设就不攻自破。已知前面的斗鸡博弈存在着纳什均衡,但不存在严格优策略均衡,这意味着纳什均衡一定是严格优策略均衡是错误的。命题得证。

再看例 2.7 上下—左中右博弈(见图 2-6)。这个博弈存在严格劣策略,但不存在严格优策略。经过重复剔除严格劣策略后,得到的均衡解为(上,中)。显然,这个均衡解也是纳什均衡,因为给定参与者 2 的策略为中,那么参与者 1 的最优策略为上;给定参与者 1 的策略为上,那么参与者 2 的最优策略为中,所以(上,中)也是纳什均衡。

命题 2.3 如果策略组合 $s^*=(s_1^*,\cdots,s_n^*)$ 为博弈 $G=\{S,u\}$ 的纳什均衡,它一定不会被重复剔除严格劣策略的方法剔除掉;如果进行了重复剔除严格劣策略后剩下唯一策略组合,那么这个策略组合一定是纳什均衡。

证:利用反证法。

如果策略组合 $s^*=(s_1^*,\cdots,s_n^*)$ 为博弈 $G=\{S,u\}$ 的纳什均衡,但却在重复剔除严格劣策略的过程中被剔除。这意味着在重复剔除严格劣策略的某个阶段,存在着某个参与者 $i\in N$,有

$$u_i(s_i^\#,s_{-i}) > u_i(s_i^*,s_{-i})$$

也就是

$$u_i(s_i^\#,s_{-i}^*) > u_i(s_i^*,s_{-i}^*)$$

成立。

显然这与 $s^*=(s_1^*,\cdots,s_n^*)$ 为博弈 $G=\{S,u\}$ 的纳什均衡相矛盾。

如果重复剔除严格劣策略后剩下唯一的策略组合不是纳什均衡，那就意味着纳什均衡在重复剔除严格劣策略的某个阶段已被剔除，显然这与前面证明的纳什均衡能够在重复剔除严格劣策略中保存下来相矛盾。命题得证。

命题 2.2 和命题 2.3 表明，纳什均衡包含着严格优策略均衡和重复剔除严格劣策略法，它是一个更为宽广、更为适用的均衡概念。

对均衡的另一种理解

在博弈论中，博弈参与者的理性、博弈结构等被处理成了公共知识，也就是它先于博弈模型而存在。但实际中，博弈参与者对博弈结构的掌握存在一个学习过程。例如，在课堂上，我们可以做一个试验，随机地选择两位同学来做囚徒困境的游戏，并不断重复。结果我们会发现，开始的时候选择沉默的次数非常多，但随着游戏的进行，选择招供的同学会越来越多，最后所有同学的选择都是招供。这说明，我们可以换一个角度来理解博弈中的均衡。例如，假设存在若干个群体(population)，我们从每一个群体中，随机地抽出一个参与者来参与博弈，对于每一次游戏，所有人都观察到博弈的整个过程和结果。虽然游戏在不断重复，但对于每一个人而言，该游戏却是一个静态游戏，因为他再次参与游戏的概率为小概率事件。由于游戏不断重复，因而最终(如果均衡存在)将会形成"一种稳定状态"，这种稳定状态就是博弈中的均衡。显然，在上述理解中，我们并不需要对博弈参与者的理性作出任何假设，但仍然得到我们所想得到的结果。这表明，博弈的理性假设是一个合理的假设。同时，理论中的博弈过程可以在试验过程中再现出来，关键在于能否对群体作出明确的区分，区分群体依赖于对偏好的定义。试验中的"稳定状态"应与理论中的均衡一致，这就意味着我们可以运用试验结果来检验理论。

第三节 简单运用举例

在这一节中将运用纳什均衡的概念来求解策略为离散情景下的简单博弈，其目的在于将前面所学内容作一个完整的总结，同时也为后面的继续学习打下

基础。

【例 2.9】 大锅饭

有两人共同抬箱子上楼，箱子抬上楼后两人会得到相应的报酬 v，由于很难判断谁出的力大，所以两人的报酬一样。如果两人同时用力，那么能够轻松地完成任务，并得到相应的报酬。如果有一个人偷懒，那么他不仅能够得到报酬，而且还不用出力，但对于用力的人而言，其收益就为报酬减去负效用（因为花费力气）。如果两人都不出力，那么箱子无法上楼，两人的收益为零。如果两人同时用力，那么负效用为 c_l，收益为 $v-c_l>0$；如果只有一个人出力，那么负效用为 c_h，收益为 $v-c_h<0$。

该博弈的基本式 $G=\{S,u\}$ 就为：

(1) 参与者集合：$N=\{1,2\}$。

(2) 策略空间：$S_1=S_2=\{出力,偷懒\}$。出力用 W 表示，偷懒用 L 表示。

(3) 偏好：每个参与者偏好为：自己偷懒，别人出力；两个人都出力；两个人都不出力；自己出力，别人偷懒。参与者 1 和 2 相应的收益函数为：

$u_1(W,W)=v-c_l \quad u_1(W,L)=v-c_h \quad u_1(L,W)=v \quad u_1(L,L)=0$

$u_2(W,W)=v-c_l \quad u_2(L,W)=v-c_h \quad u_2(W,L)=v \quad u_1(L,L)=0$

根据以上信息就可以列出相应的博弈矩阵。见图 2－10。

参与者 2

参与者 1	W	L
W	$v-c_l,v-c_l$	$v-c_h,\underline{v}$
L	$\underline{v},v-c_h$	$\underline{0},\underline{0}$

图 2－10　大锅饭博弈矩阵

这个博弈存在唯一的均衡：(L,L)。

策略组合(W,W)不是纳什均衡，因为给定参与者 2 的策略 W，参与者 1 最优的策略为 L，根据公式(2－1)，$u_1(L,W)>u_1(W,W)$。在参与者 1 对应的收益下，划一个下划线（划圆圈也行），表示给定参与者 2 的策略，参与者 1 的最优策略。

策略组合(L,W)也不是纳什均衡，因为给定参与者 1 的策略 L，参与者 2 的最优策略为 L，根据公式(2－1)，$u_2(L,L)>u_2(L,W)$。在参与者 2 对应的收益下，也划一个下划线，表示给定参与者 1 的策略，参与者 2 的最优策略。

策略组合(L,L)为纳什均衡，因为给定参与者 1 的策略 L，参与者 2 最优的策略为 L；反之，给定参与者 2 的策略 L，参与者 1 的最优策略为 L。显

然，它是唯一满足公式(2-1)的策略组合，因而参与者1和2的收益下同时有两条下划线。利用画下划线的方法可以很容易地找出两人博弈中的纳什均衡。

实际上该模型与囚徒困境完全一样，这样的模型还包括前面分析过的双头垄断、人民公社，它们都可以称为囚徒困境类博弈。大锅饭模型虽然简单，但却隐含着一个重要的经济学思想，就是所谓的道德风险。道德风险是信息经济学的重要内容，在介绍信息不完全模型时还会涉及。

㊨ 习题2.9 道德风险是指博弈的参与者故意隐藏自己的行动，而不是尽力去完成任务。通常由于没有办法有效监督参与者的偷懒行动，因而就会出现所谓的道德风险问题。在大锅饭模型中，市场是否失灵了？你是如何理解的？

【例2.10】 消耗战

有两只老虎为了争夺猎物而出现对峙，如果双方都坚持，那么就会出现僵持状态，可能双方谁都得不到猎物(猎物会腐烂)；如果有一方放弃，那么猎物就归坚持的一方所有；如果双方都放弃，那么谁也得不到猎物。猎物对老虎来说价值f，僵持的成本为c。

该博弈的基本式$G=\{S,u\}$为：

(1) 参与者集合：$N=\{1,2\}$。

(2) 策略空间：$S_1=S_2=\{P,Q\}$，P代表坚持，Q代表放弃。

(3) 收益函数：$u_1(P,P)=u_2(P,P)=-c$，$u_1(P,Q)=u_2(Q,P)=f$，$u_1=(Q,P)=u_2(P,Q)=0$，$u_1(Q,Q)=u_1(Q,Q)=0$。

相应的博弈矩阵见图2-11。

		虎2 P	虎2 Q
虎1	P	$-c,-c$	$\underline{f},\underline{0}$
	Q	$\underline{0},\underline{f}$	0,0

图2-11 消耗战博弈矩阵

该博弈存在两个纳什均衡①，分别是(P,Q)和(Q,P)。

策略组合(Q,Q)不是纳什均衡，因为给定虎2的策略为Q，那么虎1的最优策略只能是P，$u_1(P,Q)>u_1(Q,Q)$。对于虎2而言，推理相同。

策略组合(P,P)也不是纳什均衡，因为给定虎2的策略为P，那么虎1的最

① 实际上该博弈还存在一个混合纳什均衡，这涉及混合策略均衡的概念，后面会讲到。

优策略只能是 Q, $u_1(Q,P) > u_1(P,P)$。对于虎 2 而言，推理相同。

只有策略组合(P,Q)和(Q,P)才满足公式(2-1)的要求，因而是纳什均衡。消耗战模型与鹰派和鸽派、斗鸡博弈完全一样，不妨将它们称为鹰派和鸽派类博弈。这类博弈的显著特征就是博弈双方总是力求避免冲突。在政治学中，曾有人说政治就是妥协，而鹰派和鸽派类博弈就很好地描述了这一点。

【例 2.11】 性别战

有一对年轻情侣，周末他们将去听音乐会。有两场音乐会，一场为古典音乐会(用 C 表示)，另一场为流行音乐会(用 P 表示)。男的喜欢古典音乐，而女的喜欢流行音乐，但他们都不愿意分开，而是希望能够共度周末，尽管有可能参加自己所不喜欢的音乐会。

该博弈的基本式 $G = \{S, u\}$ 为：

(1) 参与者集合：$N = \{1,2\}$，男 =1，女 =2。

(2) 策略空间：$S_1 = S_2 = \{C,P\}$。

(3) 收益函数：$u_1(C,C) = u_2(P,P) = a$，$u_1(C,P) = u_1(P,C) = u_2(P,C) = u_2(C,P) = 0$，$u_1 = (P,P) = u_2(C,C) = b$，其中 $a > b > 0$。

相应的博弈矩阵见图 2-12。

		女 C	女 P
男	C	$\underline{a},\underline{b}$	0,0
	P	0,0	$\underline{b},\underline{a}$

图 2-12 性别战博弈矩阵

根据公式(2-1)，可知策略组合(C,C)和(P,P)为纳什均衡。

策略组合(C,P)不是纳什均衡，因为给定女的策略为 P，那么男的最优策略是 P(而不是 C)，$u_1(P,P) = b > 0 = u_1(C,P)$。

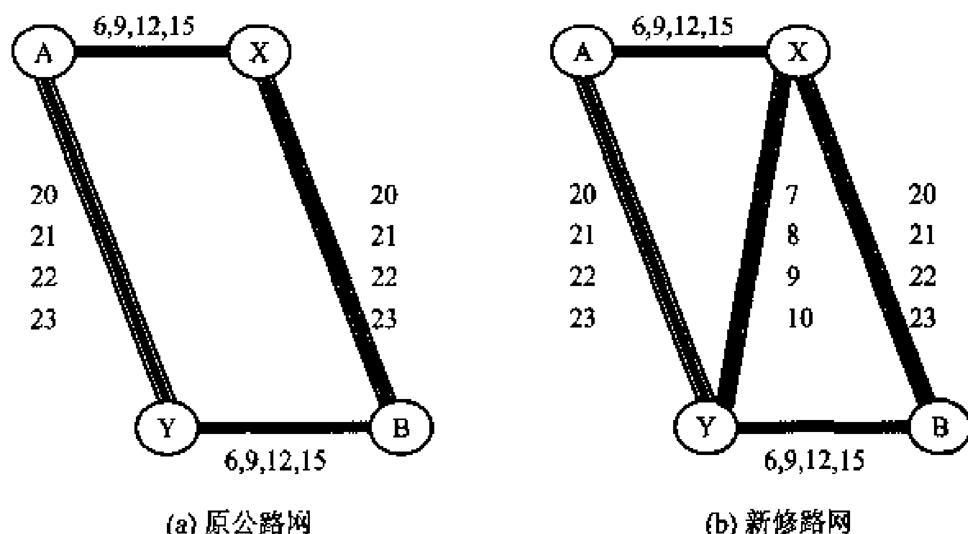
策略组合(P,C)同样不是纳什均衡，因为给定女的策略为 C，那么男的最优策略是 C(而不是 P)，$u_1(C,C) = a > 0 = u_1(P,C)$。

猎鹿，烛光晚餐和性别战这类模型实际上是完全一样的，这类模型统称为协调类模型。

【例 2.12】 最优路径

有 4 人在同一时间必须驾车从北京东站赶到北京飞机场，有两条道路可供选择，一条道路经过 X，称为 X 路，另一条道路经过 Y，称为 Y 路。从北京东站到 X 和从 Y 到北京飞机场都为小巷，只有一辆车时，需要开车 6 分钟，

每当增加一辆车时开车时间将延长3分钟(因为拥挤)。从X到北京飞机场和北京东站到Y都是高速路,只有一辆车时,需要开车20分钟,每当增加一辆车时,X到飞机场开车时间将延长1分钟,而东站到Y开车时间也将延长1分钟。见图2-13(a),A为北京东站,B为北京飞机场,数字为相应开车时间。



(a) 原公路网　　(b) 新修路网

图2-13　最优路径

该博弈的基本式 $G=\{S,u\}$ 为:

(1) 参与者集合:$N=\{1,2,3,4\}$。

(2) 策略空间:$S_i=\{X,Y\}, i\in N$。

(3) 偏好:花费的时间越少越好。

由于参与者超过了两个,所以无法画出矩阵图,但并不影响我们进行分析。求解该博弈的一个有效思路就是首先"猜"出可能的纳什均衡,然后再加以证明。这种方法对于参与者超过2人的复杂博弈特别适用。

首先,最合理的猜测是每条路线各有两辆车是一个纳什均衡。随后的关键是,证明在这种状态下,第 i 个参与者的策略会不会改变。不妨假设两人走X路,1人走Y路,那么对于第 i 个参与者来说,如果他走Y路的话是不是最优的?显然,如果他改走X路,那么花费的时间为 $12+22=34$,远远大于他走Y路的时间 $9+21=30$。所以给定其他参与者的选择,他走Y路是最优的。对于走X路的参与者而言,情况也是如此。如果两人走Y路,1人走X路,现在第 i 个参与者不走X路而改走Y路,那么花费的时间为 $22+12=34$,远远大于走X路的时间 $9+21=30$。因而,3人走同一条路,1人走另一条路不是纳什均衡,4人同走一条路也不是纳什均衡,"唯一"的纳什均衡为每条路各有两人。实际上,这个博弈存在着6个纳什均衡。

现在在 X 和 Y 之间修建了一条新的高速路，开车需要 7 分钟，每增加一辆车将增加 1 分钟，见图 2－13(b)。

㊥ 习题 2.10 新公路修建之后，新的纳什均衡是什么？会出现什么样的结果？如果我们考虑的不是汽车，而是车流，例如，存在 4 条车流，如果一条车流走 A—X—B，两条车流走 A—Y—B，最后一条车流走 A—X—Y—B，那么在 A—X 段花费 9 分钟(因为存在两条车流)，在 X—Y 段花费 7 分钟(因为存在一条车流)，在 Y—B 段花费 12 分钟(因为存在三条车流)，那么对于车流而言，纳什均衡又是什么？与不修新路相比会造成什么样的福利损失？

在上面介绍的所有模型中，纳什均衡都是一个严格均衡，因为在给定其他参与者的策略不变的情况下，任一参与者的均衡策略所得收益要严格大于其他策略所得收益。例如，在消耗战模型中，纳什均衡(P,Q)和(Q,P)是严格均衡，因为给定其他参与者的策略(不妨假设为 Q)，任一参与者的均衡策略 P 所得收益为 f，要严格大于 0；如果给定其他参与者的策略为 P，那么均衡策略 Q 所得收益为 0，要严格大于 $-c$。精确地说，就是当且仅当公式(2－1)对所有参与者而言严格不等式成立时，一个纳什均衡为**严格均衡**，否则就为**非严格均衡**。

【例 2.13】 公共产品的提供

一个小镇有 n 个居民，从小镇到外面需要经过一条河流，因而需要建设一座桥。建设桥的费用需要通过向居民征收人头税的方式来筹集。不妨设人头税为一个固定数 t，缴不缴税完全自愿。如果有大于等于 k 个居民缴税，那么桥能够建设起来，如果缴税居民数小于 k，那么由于费用不够桥将无法建设起来。不妨设 $2 \leqslant k \leqslant n$。桥对于第 i 个居民而言，价值 $v_i > t > 0$。

该博弈的基本式 $G = \{S, u\}$ 为：

(1) 参与者集合：该小镇的所有居民，$N = \{1, \cdots, n\}$。

(2) 策略空间：$S_i = \{T, E\}$，$i \in N$。T 表示纳税，而 E 表示逃税。

(3) 偏好：对于每一个小镇居民而言，最好的结果是桥建起而税没有交，收益函数 $u_i = v_i$；其次是桥建起也交了税，收益函数 $u_i = v_i - t$；较差的结果是桥没有建起也没有交税，收益函数 $u_i = 0$；最差的结果是桥没有建起而自己却交了税，收益函数 $u_i = -t$。

我们仍然用先猜后证的方法来寻找纳什均衡。

首先，考虑 $k = n$ 的情况。

当 $k = n$ 时，公共产品的提供就变成了协调类博弈，它与猎鹿模型完全一样。纳什均衡实际上只有两个：要么每一个居民都缴税，要么每一个居民都逃税。如果给定 $k-1$ 个居民都缴税，那么第 i 个居民最优的策略只能缴税，因为缴税收

益为 $v_i - t > 0$，比逃税（逃税将使桥无法建起）要好。如果 $k-1$ 个居民都逃税，那么第 i 个居民的最优策略只能是逃税，因为逃税收益为 $0 > -t$，比缴税（即使缴税桥也建不起来）要好。所以（每个居民都缴税）和（每个居民都逃税）就是纳什均衡。

其次，考虑 $k < n$ 的情况。

当 $k < n$ 时，存在着搭便车的情况，即有的居民逃税而桥仍然能够建起来。不妨猜测（k 个居民缴税，而 $n-k$ 个居民逃税）是一个纳什均衡。其证明分两种情况：一种是 $k-1$ 个居民缴税，而 $n+1-k$ 个居民逃税，那么第 i 个居民的最优策略是什么？显然，如果他逃税，桥无法建起，其收益为 0，要小于他缴税，桥建起的收益 $v_i - t$，所以第 i 个参与者缴税是最优策略。另一种情况是，已经存在 k 个居民缴税，而 $n-k-1$ 个居民逃税，那么第 i 个居民的最优策略是什么？显然，他缴不缴税都不影响桥的建设，但缴税却要付出 t，所以逃税是最优策略。因而，（k 个居民缴税，$n-k$ 个居民逃税）就是纳什均衡，其数目为 C_n^k（数学组合）。例如 $n=4, k=3$，那么纳什均衡就有 4 个。有了上面的分析，就很容易知道缴税的居民数小于 $k-1$ 或者大于 k，都不是纳什均衡，因为都存在居民会改变其策略。最后，就得到了最后一个纳什均衡，即所有居民都逃税。总结上述分析，得如下结论：

当 $k=n$ 时，（所有人纳税）或（所有人逃税）是纳什均衡。

当 $k<n$ 时，（所有人逃税）或（k 个居民纳税，$n-k$ 个居民逃税）是纳什均衡，后者的数目为 C_n^k。

显然，公共产品的提供模型中的纳什均衡都是严格纳什均衡。

(易) 习题 2.11　请你证明为什么说缴税的居民数小于 $k-1$ 或者大于 k，都不是纳什均衡。

【例 2.14】 投票

在前面曾介绍过投票模型，并指出其中一个纳什均衡是所有的村民都投票给一个候选人，而这个纳什均衡包含着劣策略。现在我们就找出投票模型中的所有纳什均衡。

在某个乡村，人数为奇数 k 的村民投票选举该村的村长，每人一票，获得多数选票的候选人获得村长职务。村长候选人有两位：A 和 B。有的村民喜欢 A，而另一些村民喜欢 B，自己喜欢的候选人获胜要好过自己不喜欢的候选人获胜。由于村民数为奇数，因而不会出现两个候选人势均力敌的情况。假设投票成本为 0，自己喜欢的候选人获胜收益为 $v > 0$，自己不喜欢的候选人获胜收益为 0。

其博弈的基本式 $G = \{S, u\}$ 为：

(1) 参与者集合：$N=\{1,\cdots,k\}$，$k=2n+1$，n 为自然数。

(2) 策略空间：$S_i=\{A,B\}$，$i\in N$。

(3) 偏好：对于每个村民而言，自己喜欢的候选人获胜是最好的，不好的就是自己不喜欢的候选人获胜。自己喜欢的候选人获胜，收益函数 $u_i=v>0$，自己不喜欢的候选人获胜，收益函数 $u_i=0$。

首先，猜测所有村民投票给一个候选人是纳什均衡。如果 $k-1$ 个村民投票给了 A(或是 B)，那么第 i 个村民无论投票给 A 还是 B，都不能改变选举结果，因而其收益是一样的，显然满足公式(2-1)的要求，因而所有村民投票给一个候选人是纳什均衡。

其次，每一个村民投票给自己喜欢的候选人也是纳什均衡。显然，每一个村民投票给自己喜欢的候选人是一个优策略，因为给定其他村民的投票，如果他投票给不喜欢的候选人，有可能出现两种结局：一种是不改变投票结局，另一种是帮助不喜欢的候选人获胜，显然这比投票给自己喜欢的候选人所获收益要小($0<v$)。因而，给定其他村民的投票，第 i 个村民最优的策略是投给自己喜欢的候选人。所以，每一个村民投票给自己喜欢的候选人为纳什均衡。

当 $k=3$ 时，上面所提到的纳什均衡即为所有的纳什均衡。

当 $k>3$ 时，分析就会变得复杂一些，这时需要考虑两种情况。

第一种情况是，一个候选人(不妨设 A)得到的选票数大于等于$\frac{k-1}{2}+2$。在这种情况下，无论第 i 个村民如何投票都不会改变投票结果(原来投 A，现在改投 B，A 仍然领先；或是原来投 B，现在改投 A，使得 A 的优势更大)，因而当一个候选人获得的票数大于等于$\frac{k-1}{2}+2$ 时，相关的所有策略组合都是纳什均衡。实际上这种情况可能包含着上面提到的两种纳什均衡，例如，所有村民都投票给 A；喜欢 A 的村民大于等于$\frac{k-1}{2}+2$，他们都投票给 A。

现在考虑第二种情况，一个候选人(不妨设 A)得到的选票数等于$\frac{k-1}{2}+1$，那么第 i 个村民如何投票就要看情况。如果每一个村民都把票投给了自己喜欢的候选人，那么第 i 个村民如果改变投票，将不会使自己的收益变好。因为，如果他喜欢 A，并且投了 A 的票，而现在改变投票，投给 B，就会使得 B 获胜，这显然要比投 A 要差。如果他喜欢 B，并投票给 B，而现在改变投票，投给 A，那么结局不变。因而，在这种情况下，纳什均衡只能是每一个村民都投票给自己喜欢的候选人。如果存在着村民把票投给了自己不喜欢的候选人这种情况，那么第 i

个村民就有积极性改变投票，因为改变投票能使自己喜欢的人获胜，因而其不可能是纳什均衡。

总结以上内容就得到投票博弈的所有纳什均衡：① 任何情况下，所有村民都投票给一个候选人是一个纳什均衡。② 任何情况下，每一个村民都投票给自己喜欢的候选人同样是一个纳什均衡。③ 当 $k>3$ 时，如果一个候选人获得的票数大于等于 $(k-1)/2+2$，那么相对应的所有策略组合都是纳什均衡。

这里我们看到，所有村民都投票给一个候选人是一个非严格纳什均衡。

㊥ *习题 2.12　如果村民不是奇数而是自然数，那么纳什均衡又是什么？*

第四节　最优反应函数

对于策略是离散的博弈，可以通过穷举策略组合，并检查其是否满足定义 2.6 来找到纳什均衡。但对复杂一点的博弈，例如策略是连续的，这种方法就难以实行。这时，一个更为一般的方法是运用最优反应函数（或对应）来寻找纳什均衡[①]。

在第一章例 1.1 石头、剪刀、布中，为了说明孙悟空和猪八戒的策略具有相互依存性，我们曾引人最优反应函数 B_s（孙悟空的最优反应函数）和 B_z（猪八戒的最优反应函数），用于描述孙悟空和猪八戒针对对手策略所作出的最优反应。例如 B_s（石头）= 布，即如果猪八戒出石头，孙悟空的最优策略为出布。这种最优反应体现的正是博弈中参与者策略之间的相互依存性，因而只要是博弈，最优反应函数就应当是存在的。

最优反应函数正式的定义如下：

定义 2.7　将映射 $B_i(i\in N)$ 称为**最优反应函数（对应）**，如果

$$B_i(s_{-i}) = \{s_i^* \in S_i : u_i(s_i^*, s_{-i}) \geq u_i(s_i, s_{-i}), \forall s_i \in S_i\}$$

上述定义中的不等式实际上就是公式（2－1），由于有可能出现等号，所以 s_i^* 有可能不唯一，在策略为连续的情况下，甚至可能存在无穷多个最优策略。为了不至于混淆，以后 B_i 就表示最优反应对应（多个），而 b_i 就表示最优反应函数（只有一个值）。

实际上，定义 2.7 与定义 2.6 完全等价。

---

① 所谓函数是指单值映射 $y=f(x)$，即任意一个 x 都有唯一的 y 与之对应。对应则是指点集映射 $Y=f(x)$，即任意一个 x 都有一个集合 Y 与之对应。显然，如果每个集合 Y 只有一个元素 y，对应就变成函数。

命题 2.4　在博弈 $G=\{S,u\}$ 中，策略组合 $s^*=(s_i^*,s_{-i}^*)$ 为纳什均衡，如果

$$s_i^* \in B_i(s_{-i}^*),\quad \forall i \in N。\qquad (2-2)$$

㊥ 习题 2.13　请你证明命题 2.4。

【例】　消耗战

根据前面的说明，我们已经知道消耗战存在两个纳什均衡：(P，Q) 和 (Q，P)。现在运用最优反应函数来求解消耗战的纳什均衡。将虎 1 的策略写在纵轴，将虎 2 的策略写在横轴。当虎 2 选择 P(坚持)时，虎 1 的最优反应策略为 Q(放弃)，因而最优反应函数为

$$b_1(P)=Q$$

当虎 2 选择 Q 时，虎 1 的最优反应策略为 P，其最优反应函数为

$$b_1(Q)=P$$

当虎 1 选择 P 时，那么虎 2 的最优反应函数为

$$b_2(P)=Q$$

当虎 1 选择 Q 时，虎 2 的最优反应函数为

$$b_2(Q)=P$$

用空心圆来表示虎 1 的反应函数，用实心圆来表示虎 2 的反应函数，最终得到图 2－14。

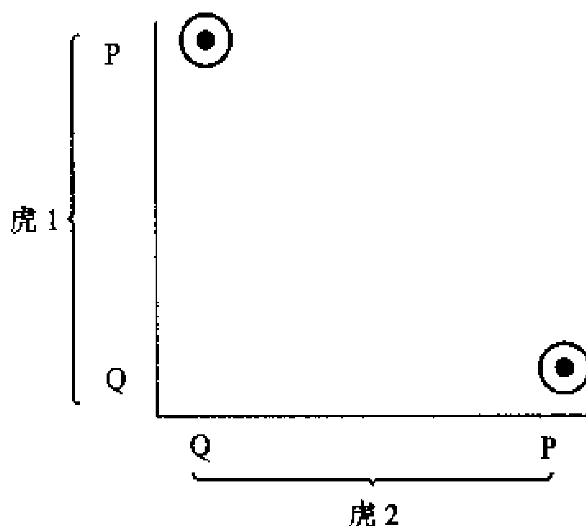


图 2－14　消耗战的最优反应函数

当空心圆和实心圆重叠时，表明(P，Q)和(Q，P)满足公式(2－2)，即它们是消耗战中的纳什均衡，与前面的分析完全一致。

【例】　石头、剪刀、布

现在用最优反应函数来看看，石头、剪刀、布这个模型是否存在纳什均衡。

首先，孙悟空的最优反应函数为

$$b_1(s_2)=\begin{cases}\text{布} & s_2=\text{石头}\\ \text{石头} & s_2=\text{剪刀}\\ \text{剪刀} & s_2=\text{布}\end{cases}$$

猪八戒的最优反应函数为

$$b_2(s_1)=\begin{cases}\text{布} & s_1=\text{石头}\\ \text{石头} & s_1=\text{剪刀}\\ \text{剪刀} & s_1=\text{布}\end{cases}$$

其次，在坐标图上画出相应的最优反应函数，孙悟空的最优反映函数用空心圆代表，猪八戒的最优反应函数则用实心圆代表，见图 2－15。从图 2－15 可以看出，没有任何一个空心圆与实心圆重叠，这表明石头、剪刀、布博弈中没有任何一个策略组合满足公式(2－2)，因而“不存在”纳什均衡。

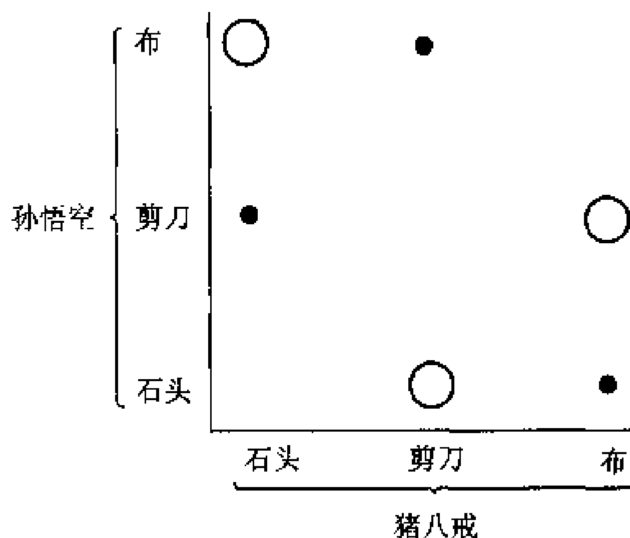


图 2－15　石头、剪刀、布的最优反应函数

【例 2.15】　协同作业

一项任务需要两人共同努力来完成。具体来讲，给定第 j 个人的努力，第 i 个人的收益随着他的努力水平先增加后下降。特别地，努力水平 $a_i \geqslant 0$，$i=1,2$。两人的收益函数如下

$$u_i(a_i)=a_i(c+a_j-a_i)\quad i,j=1,2;i\neq j$$

其中，a_i 为 i 的努力水平，a_j 为 j 的努力水平，$c>0$ 为常数。

该博弈的基本式 $G=\{S,u\}$ 为：

(1) 参与者集合：$N=\{1,2\}$。

(2) 策略空间：$S_i=[0,\bar{e}]$，$i\in N$。$\bar{e}$ 表示努力水平有一个上限，不可能无穷大，就像一天最多只有 24 小时一样。策略为努力水平 $\bar{e}\geqslant a_i\geqslant 0$。

(3) 偏好：两个参与者的偏好由其收益函数表示

$$u_i(a_i)=a_i(c+a_j-a_i)\quad i,j=1,2;i\neq j$$

由于每个参与者的策略都是连续的，因而不可能像前面那样，运用类似于离散的方法来寻找纳什均衡。这时，运用最优反应函数成为寻找该博弈纳什均衡最为有效的(有些情况下甚至是唯一可行的)方法。

给定 a_j，参与者 i 的收益函数实际上是一个二项式函数，当 $u_i(a_i)=0$ 时，其解一个是 $a_i=0$，一个是 $a_i=c+a_j$。收益函数的形状如图 2－16 所示。

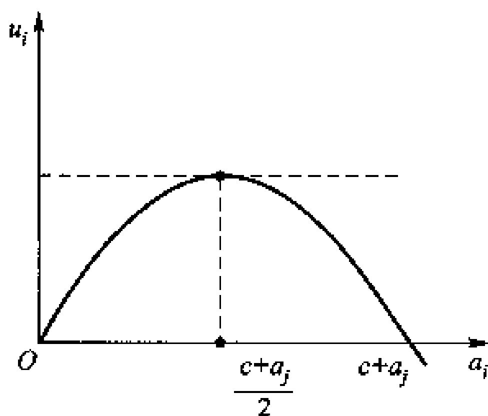


图 2－16　收益函数

根据收益函数的形状可知其一定存在最大值，当 $a_i=(c+a_j)/2$ 时，收益函数的值最大。由于参与者总是寻求收益最大化，这就意味着对于第 i 个参与者而言，给定某个 a_j 值，总存在一个唯一的 a_i 使得参与者 i 的收益最大，即存在最优反应函数

$$b_i(a_j)=\frac{1}{2}(c+a_j)$$

根据命题 2.4，可知纳什均衡就是同时满足下列联立方程的解。

$$\begin{cases}b_1(a_2)=a_1=\dfrac{1}{2}(c+a_2) & (1)\\[2ex] b_2(a_1)=a_2=\dfrac{1}{2}(c+a_1) & (2)\end{cases}$$

将(1)式代入(2)，整理后得

$$a_1=c$$
$$a_2=c$$

所以，协作博弈的纳什均衡为($a_1=c,a_2=c$)。这一结果也可以通过图 2－17来加以说明。在图 2－17 中，横轴代表 a_1，纵轴代表 a_2，两条直线分别代表着最优反应函数。当 $a_2=0$ 时，第一个参与者的最优策略为 $c/2$。当 $a_2=c$时，第一个参与者的最优策略为 c。两点决定一条直线，即得到第一个参与者的最优反应函数(直线)$b_1(a_2)$。同理可得第二个参与者的最优反应函数 $b_2(a_1)$。当两条最优反应函数相交时，其交点($a_1=c,a_2=c$)就是纳什均衡。

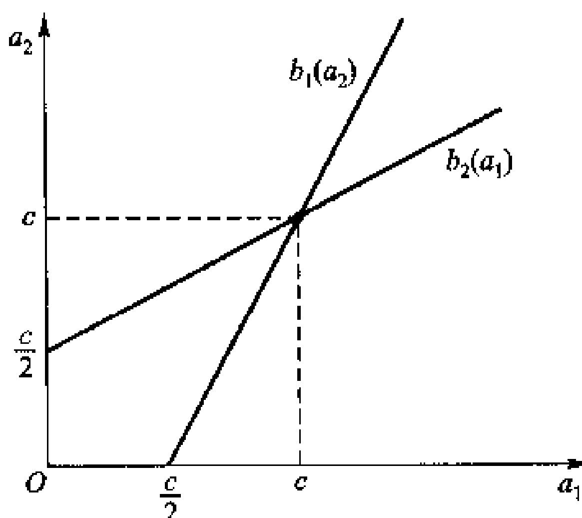


图 2-17　最优反应函数与纳什均衡

【例 2.16】　桥

在云南的山区，有一个美丽的村庄，村庄与外界唯一的联系是一座悬浮桥，为了便于说明，我们假设村庄只居住着两户人家（实际上存在 n 户人家分析也一样）。显然，悬浮桥的正常运转需要两户人家共同维护。维护需要耗费一定的成本 c_i，但同时也能得到相应的效用 $v_i(c_1+c_2)$，$i=1,2$。成本可以是时间、金钱、精力、劳动等，效用则为桥所带来的交通便利和与外界的联系，其满足边际效用递减规律。两户人家对公共产品——桥的收益就为

$$u_i(c_1,c_2) = v_i(c_1+c_2) - c_i \quad i = 1,2$$

该博弈的基本式 $G=\{S,u\}$ 为：

（1）参与者集合：两户人家，$N=\{1,2\}$。

（2）策略集合：策略为维护费用 c_i，策略空间 $S_i=[0,\bar{c}]$，$i\in N$。$\bar{c}$ 表示维护桥的费用存在一个上限，不会是无穷大。实际上这种差别并不影响结果，因而维护费用是无穷大还是存在上限并不重要。

（3）收益函数：$u_i(c_1,c_2) = v_i(c_1+c_2) - c_i \quad i = 1,2$。

由于我们并不清楚效用函数的具体形式，因而用其他方法将无法分析该博弈。但根据边际效用递减规律，可以对收益函数的大致形状进行猜测，从而得到最优反应函数，并最终找到纳什均衡。为简便起见，着重对第一户人家进行分析。

首先，不妨假设第二户人家的维护费用 $c_2=0$，那么第一户人家的最优反应函数就为：

$$b_1(0) = c_1^* \text{ 使得 } u_1(c_1,0) \quad \text{或} \quad v_1(c_1+0) - c_1 \text{ 最大}$$

由于边际效用递减规律，可知 u_1 是先增加后下降，因而一定存在最大值，并且最大值唯一。其形状见图 2-18。

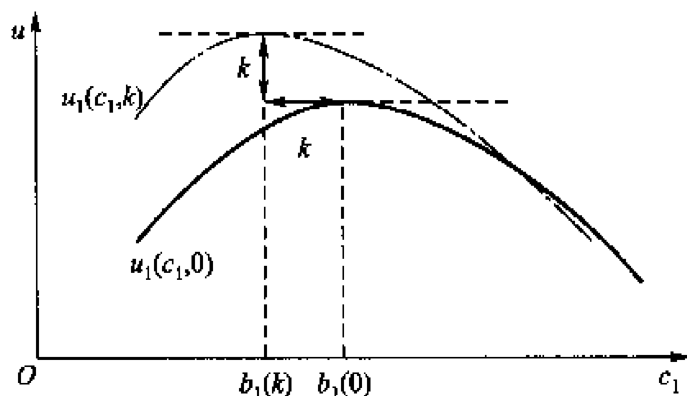


图 2－18 最优反应函数

从图 2－18 中可以看出，当 $c_2=0$ 时，由于 $v_1(c_1)$ 满足边际效用递减规律，因而随着 c_1 的增加，u_1 先增后减，所以一定存在一个 c_1^* 使得 u_1 最大。

㊥ 习题 2.14 请你证明，v_1 满足边际效用递减规律是 u_1 存在最大值的充分必要条件。

其次，考虑当 c_2 从 0 增加到 k 时的情况，这时收益函数等于

$$u_1(c_1,k)=v_1(c_1+k)-c_1=u_1(c_1+k,0)+k$$

这表明收益函数 $u_1(c_1,k)$ 的曲线实际上就是 $u_1(c_1+k,0)$ 的曲线向上移动一个数 k。由于 $c_1^*=b_1(0)$，所以 $c_1^{\#}=c_1^*-k$ 将使 $u_1(c_1,k)$ 达到最大，即 $b_1(k)=c_1^*-k$。这表明当第二户人家增加一单位维护费用，那么第一户人家的最优策略就是减少一单位费用，当第二户人家的维护费用增加到 c_1^* 时，第一户人家的最优策略（维护费用）为 0。由于可以将上面函数和变量的下标从 1 改为 2，2 改为 1（即给定第一户人家的维护费用 c_1，着重分析第二户人家），得到结果仍然一样，因而上述结论对第二户人家同样成立。由于 v_1 和 v_2 可能并不一样，因而 b_1 和 b_2 可能相等，也可能不相等。图 2－19 就反映了 $b_1(0)>b_2(0)$ 的情况。当 $b_1(0)>b_2(0)$ 时，两户人家的最优反应函数就只有一个交点，因而存在唯一的纳什均衡 $(c_1^*,0)$。当 $b_2(0)>b_1(0)$ 时，纳什均衡刚好倒过来，为 $(0,c_2^*)$。

如果 $b_1(0)=b_2(0)$，那么两户人家的最优反应曲线存在部分重合，这时只要满足 $c_1+c_2=b_1(0)=b_2(0)$，任意策略组合 (c_1,c_2) 都是纳什均衡。

在实际生活中，由于人们的偏好不尽相同，因而在公共产品的提供上唯一的纳什均衡就是一个人提供，而另一个人不提供，即存在“搭便车”行为。谁来提供关键看 c^* 的值，而与每个人的社会地位、贫富无关，既可能出现穷人搭富人的便车，也可能出现富人搭穷人的便车。

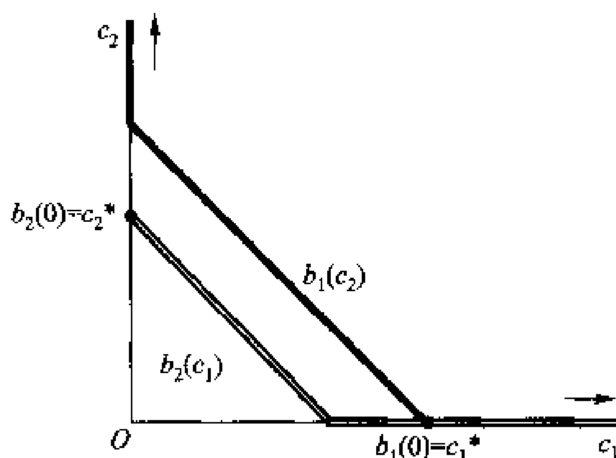


图 2-19 唯一的纳什均衡

(易) 习题 2.15 请你将前面提到的例子，例如囚徒困境、斗鸡博弈、性别战等，用最优反应函数方法求出其均衡解。

(中) 习题 2.16 大锅饭。两个工人共同完成一项工作，第 i 个工人的努力程度用 $1\geq x_i\geq 0$ 来表示，其负效用为 $c(x_i)$，$i=1,2$。这项工作完成后有收益 $R(x_1,x_2)$，不论努力程度如何，收益都将在两个工人之间平分。假设 $R(x_1,x_2)=ax_1x_2$，$a>0$。负效用分别为：(a) $c(x_i)=x_i^2$；(b) $c(x_i)=x_i$。写出该博弈的基本式，并分别求解该博弈的纳什均衡，是否存在一个策略组合比纳什均衡下的收益要高？

(中) 习题 2.17 考虑收入的公共产品提供。在例 2.16 中，参与者的财富、地位并不影响参与者对公共产品的贡献。这里我们引入收入 w_i，$i=1,2$。$w_i\geq c_i\geq 0$，为了简单化，不妨设 $w_i=w$，$i\in N$（你也可以考虑收人不相等的情况）。相应的收益函数由三部分组成，公共产品 c_1+c_2，私有物品 w_i-c_i，公共产品和私有物品的共同效用 $(w_i-c_i)(c_i+c_j)$，即收益函数变为 $u_i(c_i,c_j)=w_i+c_j+(w_i-c_i)(c_i+c_j)$，其中 $i,j=1,2$，$i\neq j$。求出该博弈的纳什均衡，是否存在一个策略组合使得所有人的收益比纳什均衡下的收益要大（即帕累托改善）？如果将参与者设定为 n，那么随着 n 的增加，纳什均衡会如何变化？是否存在帕累托改善？

【例 2.17】 古诺模型

最早对策略依存的经济现象进行研究并提出均衡解的经济学模型就数古诺模型了。古诺的论文（1838）理所当然地成为博弈论研究中的经典文献之一，并开创了产业组织理论研究的先河。后人在古诺的基础上又发展出许多变型。古诺模型的主要内容如下：

一个行业由两个寡头企业垄断，两个寡头企业生产的产品完全同质。令 q_1,q_2 分别表示企业 1 和企业 2 生产的同质产品的产量，市场中该产品的总供给为 $Q=q_1+q_2$，令 $P(Q)=a-Q$ 为需求曲线（假设如果 $Q>a$，那么 $P(Q)=0$）。

设企业 $i(i=1,2)$ 生产 q_i 的总成本 $C_i(q_i)=cq_i$，即企业不存在固定成本，且生产每单位产品的边际成本为常数 c，这里我们假定 $c<a$。

该博弈的基本式 $G=\{S,u\}$ 如下：

(1) 参与者集合：两个寡头企业，$N=\{1,2\}$。

(2) 策略集合；策略空间 $S_i=[0,a]$，策略为产量 $a\geqslant q_i\geqslant 0$，$i\in N$。

(3) 收益函数：显然寡头企业的利润函数就是垄断企业的收益函数。对于第一个企业，其收益函数 $u_1(s_1,s_2)$ 实际上为

$$\pi_1(q_1,q_2)=q_1[p(q_1+q_2)-c]$$

对于第二个企业，收益函数 $u_2(s_1,s_2)$ 为

$$\pi_2(q_1,q_2)=q_2[p(q_1+q_2)-c]$$

根据纳什均衡的定义，对于第一个企业必须满足

$$\pi_1(q_1^*,q_2^*)\geqslant\pi_1(q_1,q_2^*)$$

对所有的 $q_1\in[0,a]$ 成立。

这实际上等价于在给定第二个企业策略（产量）的情况下求第 1 个企业的利润最大化，即

$$\max_{q_1\in S_1}\pi_1(q_1,q_2^*)$$

$$\max_{0\leqslant q_1\leqslant a}\pi_1(q_1,q_2^*)=\max_{0\leqslant q_1\leqslant a}q_1[a-(q_1+q_2^*)-c]$$

显然利润函数是一个二次型函数，其形状如图 2－20（也可以对上式求一阶条件，同样能够得到一样的结果）。

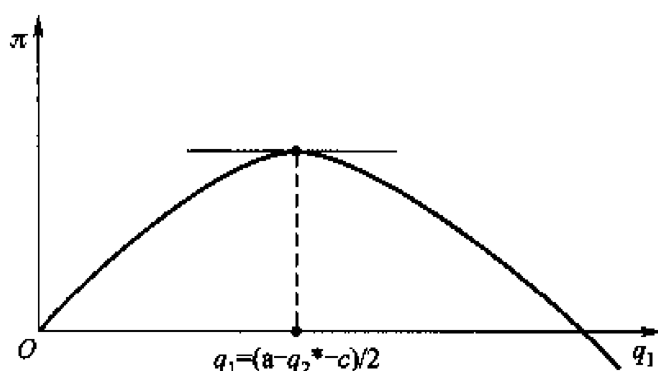


图 2－20　利润函数

当 $q_1=\frac{1}{2}(a-q_2^*-c)$ 时，利润达到最大。对于企业 2 上述分析完全一样，只需将相应的下标对换就可以了。因而纳什均衡就是下列方程组的解。

$$\begin{cases}q_1=\frac{1}{2}(a-q_2^*-c) & \text{(G-2-1)}\\ q_2=\frac{1}{2}(a-q_1^*-c) & \text{(G-2-2)}\end{cases}$$

解这组方程，得

$$q_1^* = q_2^* = \frac{a-c}{3}$$

即如果给定第二个企业的产量为 $(a-c)/3$，那么第一个企业的最优产量为 $(a-c)/3$；如果给定第一个企业的产量为 $(a-c)/3$，那么第二个企业的最优策略为 $(a-c)/3$，所以 $[(a-c)/3,(a-c)/3]$ 就为古诺模型的纳什均衡。

公式(G-2-1)和(G-2-2)即是垄断企业1和垄断企业2的最优反应函数，公式(G-2-1)和(G-2-2)构成了一个方程组，我们定义这个方程组为 $R(q_1,q_2)$，即

$$R(q_1,q_2)=\begin{cases}B_1(q_2)=\frac{1}{2}(a-c-q_2)=q_1\\ B_2(q_1)=\frac{1}{2}(a-c-q_1)=q_2\end{cases}$$

显然这个方程组是 $R: S_1\times S_2\to S_1\times S_2$ 的一个映射。它的定义域为策略组合，值域也为策略组合。通过前面的分析我们已经知道存在纳什均衡 (q_1^*,q_2^*)，即 $R(q_1^*,q_2^*)=\begin{cases}q_1^*\\ q_2^*\end{cases}$，即方程组将策略组合 (q_1^*,q_2^*) 映射成它自身，在数学中把映射成它自身的点称为不动点①。例如 $f(x)=x$，x 这个点就被称为不动点。两个寡头企业的最优反应函数也可用图2-21表示。

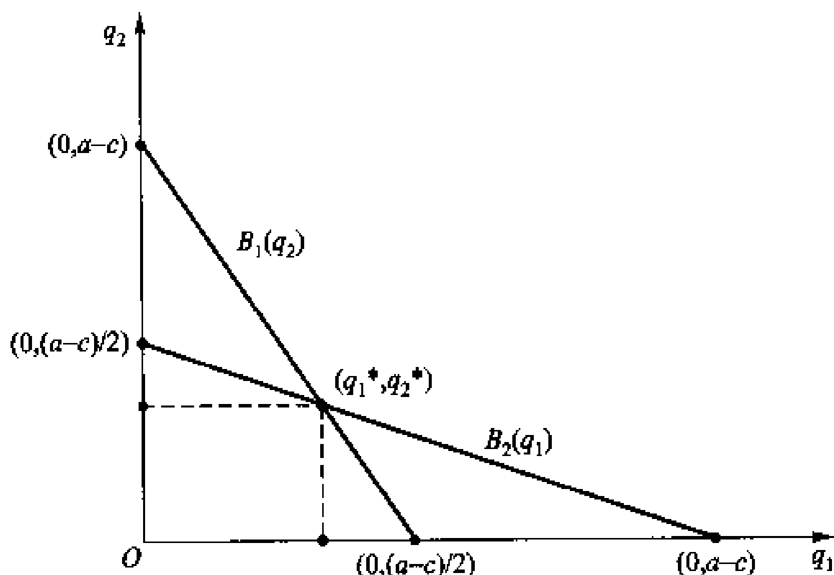


图2-21 古诺模型与不动点

① 数学中的点就是向量。

$B_1(q_2)$和$B_2(q_1)$分别为垄断企业1和2的最优反应函数,两者的交点即为纳什均衡,也是不动点。

古诺模型的意义在于指明垄断企业之间的串谋是不稳定的,每一个企业都有欺骗对方的动机。要说明这一点,假设两个垄断企业串谋在一起成立所谓的卡特尔(相当于一个企业)。根据前面的假设,这个卡特尔的最大利润是

$$\max_{0 \leq Q \leq a} \pi(Q) = pQ - cQ$$
$$\pi(Q) = Q(a - Q - c)$$

求一阶条件,得

$$Q = \frac{a-c}{2}$$

这表明如果两个垄断企业联合起来,限制总产量为$(a-c)/2$(非卡特尔下总产量为$2(a-c)/3$),就能使两个企业的总利润最大。由于两个垄断企业的成本完全一样,各自生产$\frac{a-c}{4}$的产量就能均分利润。每个企业所得利润为$(a-c)^2/8$大于非卡特尔下的利润$(a-c)^2/9$。但问题是,如果给定$q_2 = \frac{a-c}{4}$(即第二个企业遵守卡特尔协议),那么根据(G-2-1)式第一个企业的最优产量为$q_1 = \frac{3}{8}(a-c)$,显然大于卡特尔所制定的产量$\frac{a-c}{4}$。第二个企业同样也存在欺骗的动机,因而卡特尔是不稳定的,串谋双方都有动机欺骗对方。实际上,古诺模型所揭示出的卡特尔困境只不过是囚徒困境的另一种表现。

㊥ *习题2.18　如果古诺模型中的厂商数为$n>2$,那么纳什均衡下的产量如何,当n趋于无穷大时又如何?*

第五节　对称博弈和对称均衡

前面所举的大部分例子都是对称博弈,原因在于对称博弈既能说明问题,又能使我们在寻找纳什均衡的时候大大节省精力和计算。所谓对称博弈其定义如下:

定义2.8　对于策略式博弈$G=\{S,u\}$,如果每一个参与者的策略空间完全一样,即$S_i=S^*$,$\forall i \in N$,并且收益函数一致,即

$$u_i(s_i,s_j,s_{-i-j}) = u_j(s_j,s_i,s_{-i-j}) \text{ 对 } \forall i,j \in N, i \neq j \text{ 成立}$$

那么,就将策略式博弈G称为对称博弈。其中s_{-i-j}表示排除i和j之外的策略组合。

例如，在囚徒困境中，囚徒1和囚徒2的策略空间完全一样，都为招供和沉默。收益函数也是一致的，因为对于任意策略 s，$u_1(s_1,s_2)=u_2(s_2,s_1)$。比如，囚徒1选择沉默，而囚徒2选择招供，那么囚徒1所得到的收益等于囚徒1选择招供，而囚徒2选择沉默时囚徒2的收益，即

$$u_1(\text{沉默},\text{招供}) = u_2(\text{招供},\text{沉默})$$

因而，囚徒困境就是一个对称博弈。

在性别战中，虽然两个参与者的策略空间一样，但却不是对称博弈，原因在于收益函数是不一致的：$u_1(C,C)=a\neq u_2(C,C)=b$，$u_1(P,P)=b\neq u_2(P,P)=a$。

㊐ 习题2.19　请你证明石头、剪刀、布，斗鸡博弈，猎鹿，消耗战，最优路径，协同作业和桥是否为对称博弈。

通常情况下，对称博弈一般都会形成对称纳什均衡。因为参与者可以理解为来自同一个群体，他们有着相同的策略和收益函数。如果一个策略是某个参与者的最优策略，那么它也应该是另外参与者的最优策略；又由于参与者的收益函数是一致的，因而均衡下所有参与者的策略都会是一样，但也会出现例外。

定义2.9　在策略式博弈 $G=\{S,u\}$ 中，策略组合 s^* 为一个纳什均衡，如果每一个参与者的策略空间都一样，即 $S_i=S^*$，$\forall i\in N$，并且对于每一个参与者而言 s_i^* 一样，对于 $\forall i\in N$ 成立，那么 s^* 就称为对**称纳什均衡**。

显然，对称博弈不一定有对**称纳什均衡**，对称纳什均衡也不一定是对称博弈。例如，斗鸡博弈的均衡是非对称的，但它却是对称博弈；在公共产品的提供博弈中，它不是对称博弈，但却存在对称纳什均衡，所有人都纳税就是一个对称纳什均衡；而二级价格密封拍卖既不是对称博弈，也不存在对称纳什均衡。

了解对称博弈和对称均衡的好处在于可以大大简化对博弈的分析，特别是当博弈的参与者很多、参与者的策略空间很大时更是如此。遇到这种情况，我们首先判断其是不是对称博弈，如果是，将能够使分析大大简化，从而节省求解均衡的精力和时间。在后面的很多分析中将会用到这个技巧。

第六节　混合策略与纳什定理

是不是所有的完全信息静态博弈都存在纳什均衡呢？这显然是一个非常重要的问题。如果纳什均衡不存在，而我们还辛苦地寻找纳什均衡，那就是白费力气。因而，证明纳什均衡的存在性是求解博弈均衡需要首先解决的问题。根据前面给出的纳什均衡定义（见定义2.6），显然石头、剪刀、布这个博弈并不存在纳什均衡。事实真的如此吗？为了证明纳什均衡的存在性，需要对策略进行一些扩展，即将纯策略扩展为混合策略。所谓混合策略，通俗讲就是随机选择纯策

略。例如在石头、剪刀、布这个博弈中,孙悟空的一个混合策略为1/3的概率出石头、1/3的概率出剪刀、1/3的概率出布(目的是使猪八戒摸不着孙悟空的套路),记作(1/3,1/3,1/3)的形式,数学上又称为概率分布。显然混合策略可以分解为三个纯策略——石头、剪刀和布,而石头(或剪刀或布)这个纯策略不可能再分解。按照前面给出的纳什均衡定义,许多博弈是不存在纳什均衡的,例如石头、剪刀、布这个博弈。但将纯策略扩展为混合策略,纳什均衡"通常"是存在的,这就是纳什定理。下面将首先给出混合策略的概念,然后引入混合策略纳什均衡的概念,最后证明纳什定理。

一、混合策略与期望收益函数

先来看石头、剪刀、布这个博弈(见图2-22)。按照前面对纳什均衡的定义,显然该博弈不存在纳什均衡。因为每一个参与者都试图能先猜中对手可能采取的策略,从而选择相应的策略。例如,如果孙悟空猜测猪八戒可能果取石头的策略,那么孙悟空最优的策略是布,但是如果猪八戒猜测孙悟空可能采取布,猪八戒最优的策略是剪刀,而不再是石头。通俗地说,就是无论什么策略组合,孙悟空和猪八戒两者中总有一人希望改变策略,而不存在纳什均衡中的所描述的参与者策略的一致性。

		猪八戒		
		石头	剪刀	布
	石头	0,0	+1,-1	-1,+1
孙悟空	剪刀	-1,+1	0,0	+1,-1
	布	+1,-1	-1,+1	0,0

图2-22 石头、剪刀、布博弈

类似于石头、剪刀、布这样的博弈实际上是非常普遍的,例如平常的扑克游戏、体育比赛、军事对抗等。在这些博弈中,使诈是最有效的一种制胜手段,因为使诈能使对手猜不中自己要选择的策略,从面造成"不确定性"。从这一点来看,混合策略包含有一点非完全信息博弈的因素。如果参与者的最优行为是随机的,那么博弈的结果也必然是随机的。军事中的"兵不厌诈"就形象地说明了这类博弈的特点。通过引入混合策略的概念,把纳什均衡的定义扩展到包含混合策略,就可以分析诸如石头、剪刀、布这类博弈,寻找到例如扑克游戏,体育比赛,军事对抗等的均衡解。

科学地表述,纯策略是最基本的策略,它不能够分解为更小的策略;而混合策略则是参与者策略空间 S_i 中全部纯策略的概率分布,能够分解为更小的策

略。也许把纯策略(pure strategy)译为基本策略,把混合策略(mixed strategy)译为复合策略可能更为合适些。对本章所分析的完全信息静态博弈来说,一个参与者的纯策略就是他可以选择的不同行动,例如在石头、剪刀、布博弈中,策略空间 S_i 内含有三个纯策略,分别为石头、剪刀和布,而参与者 i 的一个混合策略为概率分布$(q,r,1-q-r)$,其中 q 为出石头的概率,r 为出剪刀的概率,$1-q-r\geqslant 0$ 为出布的概率,且 $0\leqslant q,r\leqslant 1$ 。由于纯策略可以看作是混合策略的一个特例,例如石头这个纯策略可以表述为混合策略(1,0,0),即以1的概率出石头,出剪刀和布的概率为0,所以混合策略包含纯策略。

作为混合策略的第二个例子,请看图2-23所示博弈。在这个博弈中,有两位参与者,每个参与者有两个纯策略:硬币的正面和背面。如果两个参与者出硬币朝上的面一样,例如参与者1出硬币的正面,参与者2也出硬币的正面,那么参与者1赢得一个硬币;如果两个参与者出硬币朝上的面不一样,例如参与者1出硬币的正面,参与者2出硬币的反面,那么参与者2赢得一个硬币。这时参与者1的一个混合策略为概率分布$(p,1-p)$,其个 p 表示出正面的概率,$1-p$ 表示出背面的概率,且 $0\leqslant p\leqslant 1$ 。在此博弈中,混合策略(1/2,1/2)表示参与者1出正面和背面的概率相同,而(1,0)表示出正面的概率为1,但绝不可能选择出背面。参与者2的情况和参与者1一样,不再赘述。参与者的纯策略只是混合策略的一种特例,混合策略(1,0)表示的就是出正面这个纯策略。

		参与者2	
		正面	背面
参与者1	正面	+1,-1	-1,+1
	背面	-1,+1	+1,-1

图2-23 猜币游戏博弈矩阵

更为一般地,假设参与者 i 有 K 个纯策略 $S_i=\{s_{i1},\cdots,s_{iK}\}$,那么参与者的一个混合策略就是一个概率分布$(p_{i1},\cdots,p_{iK})$,其中 $0\leqslant p_{ik}\leqslant 1$ 对所有 $k=1,\cdots,K$ 成立,且 $p_{i1}+\cdots+p_{iK}=1$。它表示参与者取策略 s_{i1} 的概率为 p_{i1},取 s_{iK} 的概率为 p_{iK},等。以后用 p_i 来表示基于 S_i 的任意一个混合策略,用 P_i 来表示混合策略空间(所有混合策略的集合),用 s_i 来表示任意一个纯策略。

定义2.10 对于博弈 $G=\{S,u\}$,假设 $S_i=\{s_{i1},\cdots,s_{iK}\}$。那么,参与者 i 的一个**混合策略**为映射 $p_i:S_i\rightarrow[0,1]$,即以 $p_i(s_{ik})\geqslant 0$ 的概率取 S_i 中的纯策略 s_{ik},其中 $\sum_{k=1}^{K}p_i(s_{ik})=1$ 。混合策略也可表示为 $p_i=(p_{i1},\cdots,p_{iK})$,其中对于所有的 $k=1,\cdots,K$,$0\leqslant p_{ik}\leqslant 1$,且 $p_{i1}+\cdots+p_{iK}=1$。

当参与者随机采取纯策略时，博弈的结果本身就是随机的（我们没有用不确定这个词），这时参与者的收益函数变为冯·诺伊曼—摩根斯坦恩（von Neumann - Morgenstern）收益函数，参与者 i 所得的收益就是 n 个参与者混合策略组合 $p=(p_1,p_2,\cdots,p_n)$ 下的期望值。假设 $s=(s_1,\cdots,s_n)$ 为纯策略组合，其中 s_i 表示参与者 i 可用的任意一个纯策略，$S=S_1\times\cdots\times S_n$ 为策略空间的 n 重笛卡尔积。参与者 i 从混合策略组合 p 得到的 **v - N - M 收益函数**用数学可表示为

$$v_i(p) = \sum_{s\in S} [p_1(s_1)p_2(s_2)\cdots p_n(s_n)]u_i(s) \qquad (2-3)$$

通常将这种期望偏好称为冯·诺伊曼—摩根斯坦恩（v - N - M）偏好，其中收益函数 u 称为**伯努利（Bernouli）收益函数**。对于混合策略（在微观经济学中有时又被称为博彩（lottery））p_i 和 q_i

$$p_i > q_i \text{ 当且仅当 } v_i(p_i,p_{-i}) > v_i(q_i,p_{-i}) \qquad (2-4)$$

例如，在石头、剪刀、布这个游戏中，不妨假设孙悟空有两个混合策略 (p_1,p_2,p_3) 和 (q_1,q_2,q_3)，给定猪八戒的混合策略为 $r=(r_1,r_2,r_3)$。那么对于孙悟空而言，$p>q$，当且仅当

$$\begin{aligned}
&p_1[r_1u_1(\text{石头},\text{石头}) + r_2u_1(\text{石头},\text{剪刀}) + r_3u_1(\text{石头},\text{布})] + \\
&p_2[r_1u_1(\text{剪刀},\text{石头}) + r_2u_1(\text{剪刀},\text{剪刀}) + r_3u_1(\text{剪刀},\text{布})] + \\
&p_3[r_1u_1(\text{布},\text{石头}) + r_2u_1(\text{布},\text{剪刀}) + r_3u_1(\text{布},\text{布})] > \\
&q_1[r_1u_1(\text{石头},\text{石头}) + r_2u_1(\text{石头},\text{剪刀}) + r_3u_1(\text{石头},\text{布})] + \\
&q_2[r_1u_1(\text{剪刀},\text{石头}) + r_2u_1(\text{剪刀},\text{剪刀}) + r_3u_1(\text{剪刀},\text{布})] + \\
&q_3[r_1u_1(\text{布},\text{石头}) + r_2u_1(\text{布},\text{剪刀}) + r_3u_1(\text{布},\text{布})]
\end{aligned}$$

即(2-4)式的展开。

在纯策略的情况下，理性偏好只要满足非对称性和负传递性就可以了，那么在 v - N - M 偏好下，是不是也如此呢？从公式(2-3)和(2-4)中可以看到，纯策略下的偏好是理性的并不一定意味着 v - N - M 的偏好也是理性的，除了要求满足非对称性和负传递性，还需要附加若干假设，这里我们不加证明地列出这些假设。

假设 2.1　替代公理。假设存在两个博彩（混合策略）p 和 q 以至于 $p>q$，那么对于任一 $\alpha\in(0,1)$ 和任一其他博彩 r，必有

$$\alpha p + (1-\alpha)r > \alpha q + (1-\alpha)r$$

假设 2.2　阿基米得公理。假设 $p>q>r$，那么一定存在 $\alpha,\beta\in(0,1)$，以至于 $\alpha p+(1-\alpha)r > q > \beta p+(1-\beta)r$。

相应地就有命题 2.5。

命题 2.5　定义在博彩之上的一个 v - N - M 偏好关系 $>$ 是理性的，如果满

足非对称性和负传递性,假设 2.1,假设 2.2。更进一步,如果存在一个函数 $u: S \to R$,那么

$$p_i > q_i \text{ 当且仅当 } v_i(p_i, p_{-i}) > v_i(q_i, p_{-i})$$

其中 $p = (p_i, p_{-i})$,$q = (q_i, p_{-i})$ 为混合策略组合。并且,当且仅当存在常数 $a > 0$ 和 b,以至于 $f(\cdot) = au(\cdot) + b$,那么 u 和 f 所代表的 v－N－M 偏好完全一样,即 v－N－M 偏好满足线性变换。

有了混合策略以及期望收益函数,就可以引入混合策略纳什均衡来分析具有随机性的完全信息静态博弈了。不过,在此之前先来看一看在混合策略下,严格劣策略有了什么变化。

在前面分析严格劣策略的时候,一个结论是,如果一个策略是严格劣策略,那么参与者无论在什么情况下绝对不会选择它。但人们无论什么情况下都不选择的策略是不是严格劣策略呢?在没有引入混合策略这个概念前,这个逆命题是不成立的,而引入混合策略后这个逆命题成立。我们可通过图 2－24 来加以说明。

		参与者 2	
		左	右
参与者 1	上	3,—	0,—
	中	0,—	3,—
	下	1,—	1,—

图 2－24 虚构的博弈 1

从图 2－24 中可以看出,在纯策略下参与者 1 并没有严格劣策略,因为当参与者 2 的策略为右时,参与者 1 选择下所得之收益要大于选上所得之收益(1 >0);而参与看 2 的策略为左时,参与者 1 选择下所得之收益要大于选中所得之收益(1 >0),因而不存在严格劣策略。但是,下这个策略对于参与者 1 而言,是一个从来都不会选择的策略,即便考虑混合策略也如此。

从理性出发,我们应当能够像剔除严格劣策略一样,剔除下这个策略。但在不允许混合策略的情况下,却不存在这样做的"理由"(因为它不是任何纯策略的严格劣策略)。在引入混合策略之后,我们就会看到下是一个严格劣策略。不妨假设,参与者 1 采取混合策略 $p_1 = (1/2, 1/2, 0)$,参与看 2 分别采取纯策略左和右。那么,

$$v_1(p_1, \text{左}) = 3/2 > 1 \text{ 和 } v_1(p_1, \text{右}) = 3/2 > 1$$

这意味着在允许混合策略的情况下,设 q 为参与者 2 的混合策略,那么对于任意 $q \in [0,1]$ 时,策略下绝对劣于混合策略(1/2,1/2,0),期望收益

$$v_1(p_1,q) = q_1 v_1(p_1,\text{左}) + (1-q_1)v_1(p_1,\text{右}) = 3/2 > 1$$

因而下是严格劣策略,可以剔除。相应的定义如下:

定义 2.11 在博弈 $G=\{S,u\}$ 中,p 代表混合策略,s 代表纯策略。如果对于参与者 i 而言,对于任意纯策略组合 s_{-i},始终有

$$v_i(p_i,s_{-i}) > u_i(s_i,s_{-i}) \tag{2-5}$$

就说 s_i 是参与者 i 的**严格劣策略**。这里 v_i 为 v-N-M 收益函数,u_i 为收益函数。

再看图 2-25 所示的博弈,唯一的区别在于将 1 改成了 2。

参与者 2

		左	右
参与者 1	上	3,—	0,—
	中	0,—	3,—
	下	2,—	2,—

图 2-25 虚构的博弈 2

从图 2-25 可以看出,在纯策略的情况下,策略下仍是一个从不选择的策略,但它不是严格劣策略,原因在于引入混合策略后,选择下比选择某个混合策略要优,因而其不满足"无论什么情况"这一要求。例如,设参与者 1 的混合策略为$(p_{11},p_{12},0)$,那么当 $v_1(p_1,\text{左}) > u_1(\text{下},\text{左})$时,必有

$$v_1(p_1,\text{右}) < u_1(\text{下},\text{右})$$

即(2-5)式并不是始终成立的。实际上,策略下在允许混合策略的情况下并不是从来都不选择的,例如当参与者 1 选择混合策略(1/2,1/2,0)时,其期望值为 3/2,小于 2,所以这时选择纯策略下要好过混合策略(1/2,1/2,0)。

类似地可以定义混合策略下的劣策略,即(2-5)式严格大于号变为大于等于号,但至少存在某个 s_{-i}使得严格不等号成立。

㊣ 习题 2.20 请你给出混合策略下的劣策略定义。

二、混合策略纳什均衡与纳什定理

在这一节中,分三个步骤来完成纳什定理的证明。第一步是将纳什均衡的概念扩展到混合策略,即混合策略纳什均衡;第二步求解石头、剪刀、布和猜币游戏这两个博弈的纳什均衡;第三步是一个"非数学化"的纳什均衡证明。

(一) 混合策略纳什均衡

需要记住,纯策略只不过是混合策略的特例,也就是说纯策略是混合策略的

退化,纯策略集合是混合策略集合的真子集。这就意味着对纯策略成立的要求同样对混合策略成立。

在纯策略下,纳什均衡的核心是每一个参与者所选择的策略都是对手策略的最优反应策略,并且对所有参与者都如此。在混合策略下,同样应满足这一要求(因为纯策略是混合策略的真子集)。问题是混合策略纳什均衡是不是要比纯策略纳什均衡满足更多的条件。由于混合策略反映了参与者随机地选择策略,目的就是使对手摸不清自己的套路,因而博弈的结果也带有随机性,但是收益的期望值却是确定的(正是由于这一点它被看作是完全信息,但又具有一点点非完全信息的因于)。

因而必须在两点上加以扩展:

(1) 必须在混合策略的概念下来讨论最优反应函数,也就是任意参与者的最优反应函数反映的是给定其他参与者混合策略组合下的最优混合策略。

(2) 必须根据 v-N-M 收益函数来计算收益的期望值,因而所谓最优是期望值的最优。

如果博弈为 $G=\{S,u\}$,其中 $N=\{1,2,\cdots,n\}$,$S_i=\{s_{i1},\cdots,s_{iK}\}$,$u=(u_1,u_2,\cdots,u_n)$, $\forall i\in N$,那么参与者 i 的混合策略为 $p_i=(p_{i1},p_{i2},\cdots,p_{iK})$,期望收益函数为 $v_i(p_i,p_{-i})=E[u_i(s)]$。

定义 2.12 一个混合策略组合 $p^*=(p_1^*,p_2^*,\cdots,p_n^*)$是博弈 $G=\{S,u\}$ 的一个**混合策略纳什均衡**,如果对每一个参与者 $i=1,2,\cdots,n$,对于所有的 $p_i\in P_i$,不等式

$$v_i(p_i^*,p_{-i}^*)\geqslant v_i(p_i,p_{-i}^*) \tag{2-6}$$

成立。

这就引出了一个有待证明的命题,即纯策略纳什均衡会不会在混合策略下不再是纳什均衡?

命题 2.6 纳什均衡的等价性。在不允许混合策略下得到的纯策略纳什均衡,在混合策略下仍然是纳什均衡;反之,在混合策略下,如果存在纯策略纳什均衡,那么它同样是不允许混合策略下的纯策略纳什均衡。

证明:用反证法。如果 s^* 是纯策略纳什均衡,不妨设其在混合策略下不再是纳什均衡,那么至少存在一个混合策略使得

$$v_i(p_i,s_{-i}^*)>u_i(s_i^*,s_{-i}^*)$$

不妨设参与者 i 存在 n 个行动,那么上式可分解为

$$\sum_{l=1}^{n}p_{il}u_i(s_{il},s_{-i}^*)>\sum_{l=1}^{n}p_{il}u_i(s_i^*,s_{-i}^*)$$

其中 $\sum_{l=1}^{n} p_{il} = 1$。可知 s^* 为纯策略纳什均衡，因而有 $u_i(s_i^*, s_{-i}^*) \geq u_i(s_{il}, s_{-i}^*)$，即

$$\sum_{l=1}^{n} p_{il}u_i(s_{il}, s_{-i}^*) \leq \sum_{l=1}^{n} p_{il}u_i(s_i^*, s_{-i}^*)$$

显然这是矛盾的。

反过来，如果 s^* 是混合策略下的纯策略纳什均衡，不言自明 s^* 必然是不允许混合策略下的纳什均衡。实际上纯策略空间是混合策略空间的真子集，如果 s^* 是混合策略空间中的纳什均衡，并且它属于纯策略空间，那么它也一定是纯策略空间中的纳什均衡。证毕。

显然，混合策略纳什均衡与纯策略纳什均衡并没有本质的区别，以后我们将两者统称为纳什均衡。

（二）两个例子

为了形象地理解混合策略纳什均衡，我们举两个不存在纯策略纳什均衡的例子。第一个例子为猜币游戏（见图 2－23）。这个博弈显然不存在纯策略下的纳什均衡，但却存在着混合策略纳什均衡。该博弈的基本式为：

（1）参与者：游戏者 1 定义为 1，游戏者 2 定义为 2，$N = \{1,2\}$。

（2）混合策略空间：令混合策略为 $p_i = (p_{i1}, 1 - p_{i1})$，$1 \geq p_{i1} \geq 0$ 为参与者 i 出正面的概率，$1 - p_{i1} \geq 0$ 为出背面的概率，混合策略空间为 $P_i = \{p_i : 1 \geq p_{i1} \geq 0\}$。

（3）收益函数：参与者 i 的伯努利收益函数为 $u_i(s_1, s_2)$，v－N－M 收益函数为 $v_i(p_1, p_2) = \sum_{s \in S} [p_1(s_1)p_2(s_2)]u_i(s_1, s_2)$。

给定游戏者 2 的混合策略为 $(p_{21}, 1 - p_{21})$，那么游戏者 1 出正面的期望值为

$$p_{21} \times 1 + (1 - p_{21}) \times (-1) \qquad (2-M-1)$$

游戏者 1 出背面的期望值为

$$p_{21} \times (-1) + (1 - p_{21}) \times 1 \qquad (2-M-2)$$

如果游戏者 1 的混合策略为 $(p_{11}, 1 - p_{11})$，那么它的期望收益为

$$v_1(p_1, p_2) = p_{11} \times (2-M-1) + (1 - p_{11}) \times (2-M-2)$$

整理后得

$$v_1(p_1, p_2) = (2p_{11} - 1)(2p_{21} - 1)$$

显然，当 $p_{21} > 1/2$ 时，p_{11} 越大，期望收益越大，当 $p_{11} = 1$ 时，期望收益最大。它的意思非常简单，如果游戏者 2 出正面的概率大于 1/2（出正面的可能性大于背面的可能性），游戏者 1 最优的策略反应是增加出正面的概率。如果游戏者 2 出正面的概率小子 1/2（较可能出背面），游戏者 1 最优的策略反应是增加出背

面的概率，当 $p_{11}=0$ 时，期望收益最大。显然，参与者之间的策略是相互依存的，即游戏者 1 的 p_{11} 是 p_{21} 的函数[①]，记作

$$B_1(p_{21}) = p_{11} = \begin{cases} 0 & p_{21} < 1/2 \\ [0,1] & p_{21} = 1/2 \\ 1 & p_{21} > 1/2 \end{cases} \qquad (2-M-3)$$

(2 - M - 3)式如图 2 - 26 所示。

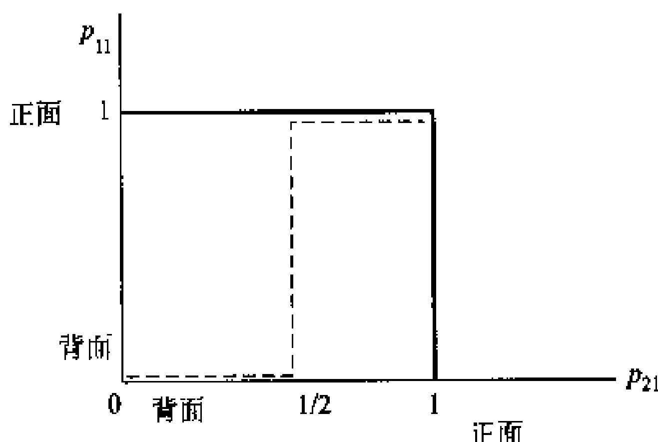


图 2 - 26 游戏者 1 的最优反应对应

给定游戏者 1 的混合策略为 $(p_{11}, 1-p_{11})$，那么游戏者 2 出正面的期望值为

$$p_{11} \times (-1) + (1-p_{11}) \times 1 \qquad (2-M-4)$$

游戏者 2 出背面的期望值为

$$p_{11} \times 1 + (1-p_{11}) \times (-1) \qquad (2-M-5)$$

如果游戏者 2 的混合策略为 $(p_{21}, 1-p_{21})$，那么他的期望收益为

$$v_2(p_1, p_2) = p_{21} \times (2-M-4) + (1-p_{21}) \times (2-M-5)$$

整理后得

$$v_2(p_1, p_2) = (1-2p_{21})(2p_{11}-1)$$

当 $p_{11}<1/2$ 时，p_{21} 越大，期望收益越大，这说明增加 p_{21} 能够增加收益。它的意思非常简单，如果游戏者 1 出正面的概率小于 1/2，游戏者 2 最优的策略反应是增加出正面的概率，当 $p_{21}=1$ 时，期望收益最大。如果游戏者 1 出正面的概率大于 1/2，游戏者 2 最优的策略反应是增加出背面的概率，当 $p_{21}=0$ 时，期望收益最大。游戏者 2 的 p_{21} 是 p_{11} 的对应，记作

$$B_2(p_{11}) = p_{21} = \begin{cases} 1 & p_{11} < 1/2 \\ [0,1] & p_{11} = 1/2 \\ 0 & p_{11} > 1/2 \end{cases} \qquad (2-M-6)$$

① 确切地说是对应。函数是点到点的映射，而对应是点到集合的映射。

(2 - M - 6)式如图 2 - 27 所示。

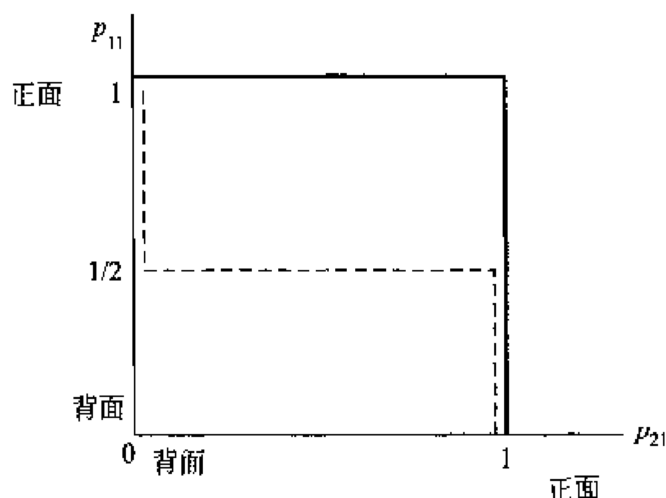


图 2 - 27 游戏者 2 的最优反应对应

如果混合策略组合(p_1^*, p_2^*)为纳什均衡,那么必须同时满足(2 - M - 3)和(2 - M - 6)两式(命题 2.4)。如图 2 - 28,只有当 $p_{11} = p_{21} = 1/2$ 时,游戏者的最优反应对应曲线才有相交点,这个交点就是纳什均衡。

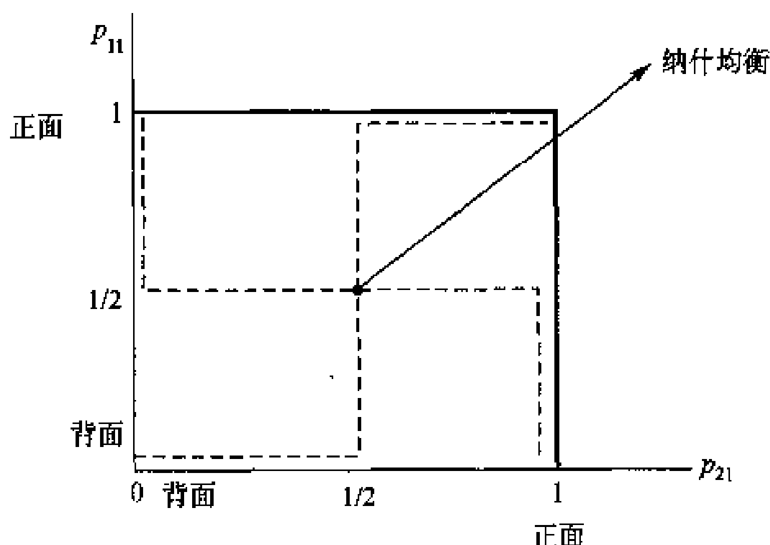


图 2 - 28 纳什均衡点

第二个例子是石头、剪刀、布(见图 2 - 22)。它比猜币博弈复杂得多。首先将这个博弈转化为基本式:

(1) 参与者:孙悟空定义为 1,猪八戒定义为 2,$N = \{1,2\}$。

(2) 混合策略空间:参与者 i 的策略空间为 $S_i = \{$石头,剪刀,布$\}$,s_i 为纯策略。混合策略为 $p_i = (p_{i1}, p_{i2}, 1 - p_{i1} - p_{i2})$,$p_{i1}$为参与者 i 出石头的概率,p_{i2}为参与者 i 出剪刀的概率,$1 - p_{i1} - p_{i2}$为出布的概率,并且满足非负与总和为 1 的条

件。混合策略空间为 $P_i = \{p_i;1 \geq p_{i1},p_{i2} \geq 0;1 - p_{i1} - p_{i2} \geq 0\}$ 。$i=1,2$。

(3) 收益函数:博弈参与者 i 的伯努利收益函数为 $u_i(s_1,s_2)$,$i=1,2$。参与者 1 的 v-N-M 收益函数为 $v_1(p_1,p_2) = \sum_{j=1}^{3} p_{1j} \sum_{k=1}^{3} p_{2k} u_1(s_{1j},s_{2k})$,参与者 2 的 v-N-M 收益函数为 $v_2(p_1,p_2) = \sum_{k=1}^{3} p_{2k} \sum_{j=1}^{3} p_{1j} u_2(s_{1j},s_{2k})$ 。

给定猪八戒的混合策略为$(p_{21},p_{22},1-p_{21}-p_{22})$,孙悟空出石头的期望值为

$$p_{21} \times 0 + p_{22} \times 1 + (1 - p_{21} - p_{22}) \times (-1) \qquad (2-S-1)$$

出剪刀的期望值为

$$p_{21} \times (-1) + p_{22} \times 0 + (1 - p_{21} - p_{22}) \times 1 \qquad (2-S-2)$$

出布的期望值为

$$p_{21} \times 1 + p_{22} \times (-1) + (1 - p_{21} - p_{22}) \times 0 \qquad (2-S-3)$$

如果孙悟空的混合策略为$(p_{11},p_{12},1-p_{11}-p_{12})$,那么它的期望收益为

$$v_1(p_1,p_2) = p_{11} \times (2-S-1) + p_{12} \times (2-S-2) + (1 - p_{11} - p_{12}) \times (2-S-3)$$

整理之后得

$$v_1(p_1,p_2) = p_{11}(3p_{22} - 1) - p_{12}(3p_{21} - 1) + p_{21} - p_{22}$$

从上式可以看出,当 $p_{22} > 1/3$ 时,p_{11}越大,期望收益越大;当 $p_{21} < 1/3$ 时,p_{12}越大,期望收益越大。通俗地说,就是如果猪八戒出剪刀的概率大于 1/3,那么孙悟空增加出石头的概率(p_{11})能够增加收益;如果猪八戒出石头的概率小于 1/3,那么孙悟空增加出剪刀的概率能够增加收益。这表明如果猪八戒的混合策略不为(1/3,1/3,1/3),那么孙悟空就会相应地增加或是减少出石头或剪刀或布的概率。另一个思考的方法就是对上式求一阶条件。考虑到运用求导的方法能够简化和统一分析,读看应该逐步适应这种方法。

对上式求一阶偏导数,并令其等于 0,得

$$\frac{dv_1}{dp_{11}} = 3p_{22} - 1 = 0 \quad \frac{dv_1}{dp_{12}} = 1 - 3p_{21} = 0$$

当 $p_{22} > 1/3$ 时, $\frac{dv_1}{dp_{11}} > 0$,这说明 p_{11}增大能够增加收益;当 $p_{21} < 1/3$ 时, $\frac{dv_1}{dp_{12}} > 0$,这说明 p_{12}增大能够增加收益。这意味着如果猪八戒的混合策略不为(1/3,1/3,1/3),那么孙悟空就会相应地增加或是减少出石头或剪刀或布的概率,与我们的直观分析一致。

现在给定孙悟空的混合策略为$(p_{11},p_{12},1-p_{11}-p_{12})$,那么猪八戒出石头的期望值为

$$p_{11} \times 0 + p_{12} \times 1 + (1 - p_{11} - p_{12}) \times (-1) \qquad (2-S-4)$$

出剪刀的期望值为

$$p_{11} \times (-1) + p_{12} \times 0 + (1 - p_{11} - p_{12}) \times 1 \qquad (2-S-5)$$

出布的期望值为

$$p_{11} \times 1 + p_{12} \times (-1) + (1 - p_{11} - p_{12}) \times 0 \qquad (2-S-6)$$

如果猪八戒的混合策略为$(p_{21}, p_{22}, 1 - p_{21} - p_{22})$,那么它的期望收益为

$$v_2(p_1, p_2) = p_{21} \times (2-S-4) + p_{22} \times (2-S-5) + (1 - p_{21} - p_{22}) \times (2-S-6)$$

整理之后得

$$v_2(p_1, p_2) = p_{21}(3p_{12} - 1) - p_{22}(3p_{11} - 1) + p_{11} - p_{12}$$

求一阶条件,得

$$\frac{\mathrm{d}v_2}{\mathrm{d}p_{21}} = 3p_{12} - 1 = 0 \qquad \frac{\mathrm{d}v_2}{\mathrm{d}p_{22}} = 1 - 3p_{11} = 0$$

当 $p_{12} > 1/3$ 时,$\frac{\mathrm{d}v_2}{\mathrm{d}p_{21}} > 0$,这说明 p_{21} 增加能够增加收益;当 $p_{11} < 1/3$ 时,$\frac{\mathrm{d}v_2}{\mathrm{d}p_{22}} > 0$,这说明 p_{22} 增加能够增加收益。通俗地说,就是如果孙悟空出剪刀的概率(p_{12})大于1/3,那么猪八戒增加出石头的概率(p_{21})能够增加收益;如果孙悟空出石头的概率小于1/3,那么猪八戒多出剪刀能够增加收益。这说明如果孙悟空的混合策略不为(1/3,1/3,1/3),那么猪八戒就会相应地增加或是减少出石头或剪刀或布的概率。

综上所述,如果猪八戒的混合策略为(1/3,1/3,1/3),那么孙悟空最优的混合策略为(1/3,1/3,1/3);反之,如果孙悟空的混合策略为(1/3,1/3,1/3),那么猪八戒最优的策略为(1/3,1/3,1/3),所以混合策略组合[(1/3,1/3,1/3),(1/3,1/3,1/3)]就为纳什均衡。

实际上,猜币游戏和石头、剪刀、布都是对称博弈,那么在混合策略下它们仍然是对称博弈吗?

定义 2.13 对于博弈 $G = \{S, u\}$,假如在纯策略下它是对称博弈,那么同样把它称为**混合策略下的对称博弈**。

同样可以定义混合策略下的对称均衡。

定义 2.14 对于博弈 $G = \{S, u\}$,混合策略组合 p 是**对称混合策略纳什均衡**,如果 p 是一个混合策略纳什均衡,并且对于每一个参与者而言混合策略都相同。

在前面曾提到,在纯策略下,对称博弈并不一定意味着对称均衡,但在混合策略下,对称博弈必然有对称均衡。

命题 2.7 对于博弈 $G=\{S,u\}$，如果每一个参与者都有相同的有限行动，那么博弈 G 一定有对称混合策略纳什均衡。

需要注意的是纯纳什均衡也是混合策略纳什均衡。

这意味着，即使将分析扩展到混合策略，对称博弈的特性仍然会得到保留，利用对称博弈的特性来简化分析仍然可行，并且均衡还是对称的。

㊣ 习题 2.21 请你证明鸽派和鹰派、性别战、斗鸡博弈存在混合策略纳什均衡，并且证明这些均衡是否为对称均衡。

为了进一步简化混合策略下的分析，有命题 2.8。

命题 2.8 对于博弈 $G=\{S,u\}$，如果每一个参与者的行动都是有限的，那么混合策略组合 p^* 为纳什均衡，当且仅当对每一个参与者 i 而言，给定其他参与者的混合策略组合 p_{-i}^*，

(1) 每一个概率为正的纯策略的期望收益都相等，即 $v_i(s_i,p_{-i}^*)=v_i(s_i',p_{-i}^*)$，其中 s_i,s_i'为 p_i^* 中概率为正的任意两个纯策略；

(2) 概率为零的纯策略的期望收益不可能大于概率为正的纯策略的期望收益，即 $v_i(s_i,p_{-i}^*)\leqslant v_i(s_i',p_{-i}^*)$，其中 s_i 为 p_i^* 中概率为 0 的任意一个纯策略，而 s_i'为 p_i^* 中概率为正的任意一个纯策略。

命题 2.7 和命题 2.8 在第三章中求混合策略纳什均衡时将会具体用到。

（三）纳什定理

在考虑混合策略的情况下，一个完全信息静态博弈“通常”都存在纳什均衡。这一认识最早由纳什（Nash，1950，1951）进行了证明，后人为表纪念，将其称为纳什定理。

定理 2.1 纳什定理。如果完全信息静态博弈 $G=\{S,u\}$ 是有限的，即参与者是有限的，S 包含的纯策略也是有限的，那么一定存在至少一个（混合策略）纳什均衡。

所谓“通常”指的就是博弈的有限性。到目前为止，我们分析的博弈都满足这两个条件。参与者有限——通常为两个人。S 也是有限的——要么存在若干个纯策略，例如石头、剪刀、布；要么虽然纯策略有无穷多，但却是有界闭集①，例如协同作业中 $S_i=[0,\bar{e}]$，桥中 $S_i=[0,\bar{c}]$，古诺模型中 $S_i=[0,a]$，所以前面分析的静态博弈都存在纳什均衡。

下面的内容涉及大量数学，读者跳过也不影响后面的阅读，只要记住纳什定理的意义即可，不过还是建议读者不妨试试。

① 数学中，有界闭集必然意味着是有限的。

对纳什定理的证明分三个步骤来完成:第一步,介绍不动点定理;第二步,说明纳什均衡就是不动点;第三步,完成证明。

第一步,不动点定理。

先来看图 2-29。

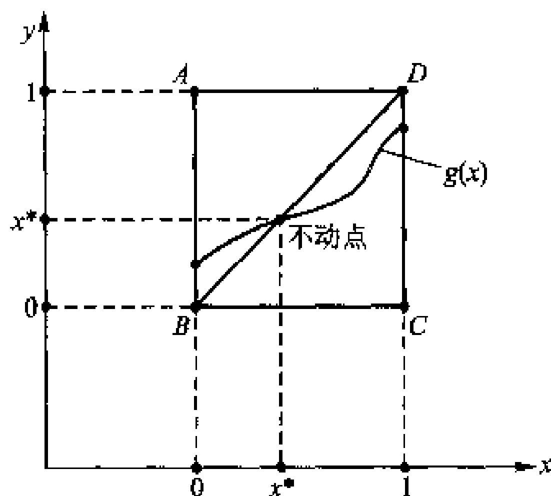


图 2-29　从区间[0,1]到区间[0,1]的连续函数具有不动点

图 2-29 中的方框为正方形。BD 为 45°线,它实际上等价于函数 $y=f(x)=x$。BD 线上的每一点都满足 $f(x)=x$,把满足 $f(x)=x$ 的 x 称为**不动点**。BD 线所代表的函数的自变量 x 取值为[0,1],称为定义域,而因变量 y 的取值为[0,1],称为值域。定义域和值域都为[0,1]的函数 $f(x)$,又可写作 $f:[0,1]\to[0,1]$,读作从[0,1]到[0,1]的映射。现在做一个小游戏:拿一支笔,从 AB 线出发任意画一条曲线,记住可以任意画,但必须连续画不能断裂,直到 CD 线结束。这时你会发现无论如何画,这条曲线一定会和 BD 线相交,这个交点就是不动点。因为任意画的曲线等价于一个函数 $g(x)$,没有断裂意味着它是连续的,与 BD 线相交意味着 $g(x^*)=x^*$,所以交点为不动点。

所有任意画的曲线都有如下的要点:

(1) 定义域[0,1]是非空的、有界的、闭的、凸集合,简称非空凸紧集。有界闭集即为紧集。

(2) 函数 $f(x)$ 是连续的,并且是自身对自身的映射。

上面考虑的只是一元函数,对于 n 元函数是不是如果满足上述要点就存在着不动点呢?数学家布劳尔(Brower)证明了这一点。

定理 2.2　布劳尔(Brower)不动点定理。假设 $A\subset R^n$ 是一个非空凸紧集合,并且 $f:A\to A$ 是一个连续函数;那么 $f(\cdot)$ 必有一个不动点;即,存在一个 $x\in A$ 使得 $x=f(x)$。

这里 R 为实数,R^n 为实数的 n 重笛卡尔积,x 为 n 维向量。在前面的例子中

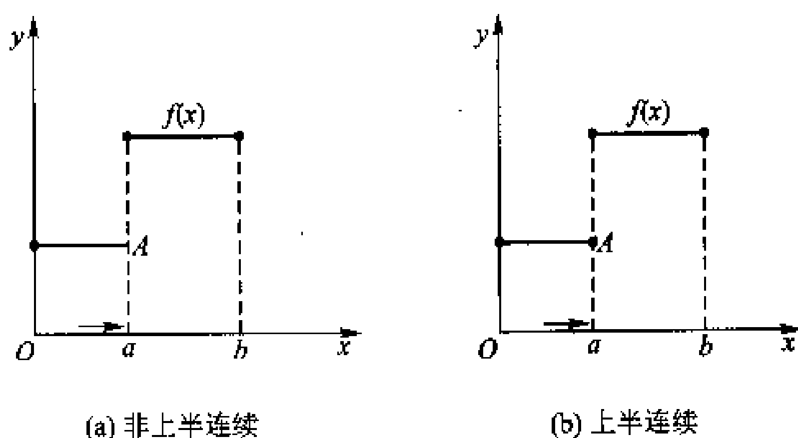
$n=1, A=[0,1]$。在实际运用中,我们常常遇到的不是函数而是对应的情况,例如猜币游戏这个博弈存在的是最优反应对应,而不是函数,这就有了对布劳尔不动点定理的推广。

定理 2.3　角谷(Kakutani)不动点定理。 假设 $A \subset R^N$ 是非空凸紧集,并且 $f: A \to A$ 是上半连续对应,并且对于每一个 $x \in A$,集合 $f(x) \subset A$ 是非空凸集,那么 $f(\cdot)$ 必有一个不动点。即,存在一个 $x \in A$ 使得 $x \in f(x)$。

角谷不动点定理与布劳尔不动点定理的区别在于:① 映射 f 为对应而不是函数,所以 $f(x)$ 为一个集合(一个点也能组成集合,所以函数是对应的特例)。② 映射 f 是上半连续的。一个对应只有既是上半连续又是下半连续才能称为连续,因而上半连续比连续的要求要弱。

关于对应的**上半连续**,通俗讲就是如果存在一个序列 $x^1, x^2, \cdots, x^n$ 收敛于 x,并且 $y^1 \in f(x^1), y^2 \in f(x^2), \cdots, y^n \in f(x^n)$ 收敛于 y,那么一定有 $y \in f(x)$(如果 $f(x)$ 只有唯一值,那么 $\in$ 就变为 $=$),即对应包含着它的极限点。

在图 2-30(a)中,当 $x \to a$,即从 a 的左边向其收敛时,$f(x) \to A$,但我们看到 $A \notin f(a)$,即对应 $f(x)$ 并不包含极限点 A,所以它不是上半连续的。而在图 2-30(b)中,对应 $f(x)$ 包含它的每一个序列的极限点,特别是 $A \in f(a)$,因而它满足上半连续,但它不是连续的。例如,猜币游戏中每个参与者的最优反应对应就是上半连续的。对于函数而言,上半连续就等同于连续。



(a) 非上半连续　　(b) 上半连续

图 2-30　对应的上半连续

第二步,纳什均衡是不动点。

根据纳什均衡的定义,在给定其他参与者策略的情况下,参与者 i 选择一个策略以使其收益最大化,并且对每一个参与者都如此,即纳什均衡是一个使每一个参与者收益最大的策略组合。针对其他参与者不同的策略,参与者 i 都有一个最优策略与之相对应,用数学来表示就是最优反应函数或对应。如果参与者 i

的最优反应函数(对应)定义为 $B_i(p_1,\cdots,p_{i-1},p_{i+1},\cdots,p_n)=p_i$,其中 $i=1,\cdots,n$,那么 n 个参与者的最优反应函数(对应)就构成了一个方程,将其定义为 $R(p_1,\cdots,p_n)$,即

$$R(p_1,\cdots,p_n)=\begin{cases}B(p_2,\cdots,p_n)=p_1\\ \vdots\\ B_i(p_1,\cdots,p_{i-1},p_{i+1},\cdots p_n)=p_i\\ \vdots\\ B_n(p_1,\cdots,p_{n-1})=p_n\end{cases}$$

如果 $(p_1^*,\cdots,p_n^*)$ 是纳什均衡,它一定满足 $R(p_1^*,\cdots,p_n^*)=(p_1^*,\cdots,p_n^*)$,即纳什均衡是一个不动点。例如,前面分析的古诺模型中的最优反应函数(G-2-1)和(G-2-2)共同构成了一个方程 $R(q_1,q_2)$,这个方程的解既是不动点,也是纳什均衡。

第三步,完成证明。

现在要完成的是证明:① 最优反应对应 B_i 的定义域为非空凸紧集;② B_i 为上半连续的对应。由于 R 是 $B_i(i=1,\cdots,n)$ 的方程组,B_i 满足这两个条件,总的策略方程 R 必然满足这两个条件,而根据不动点定理,我们知道存在着不动点 $(p_1^*,\cdots,p_n^*)$,使得 $R(p_1^*,\cdots,p_n^*)=(p_1^*,\cdots,p_n^*)$。

(1) 定义域为非空凸紧集。所谓有界集合就是任意画一个圆(圆的半径可以任意大,但不能无穷大)可以把它包下。所谓闭集就是包含边界的集合。有界闭集即为紧致集合,简称**紧集**。所谓凸集就是在集合内任意找两点,连接两点的直线一定位于集合中。图 2-31 中的直线和三角平面显然是有界的,同时也是闭的,即它们是紧集。同时它们也是凸集。

而图 2-32 中的环是非空紧集,但却不是凸集。

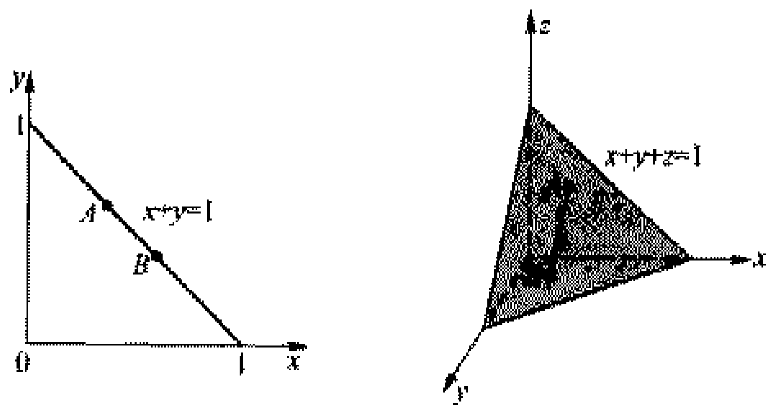


图 2-31 非空凸紧集

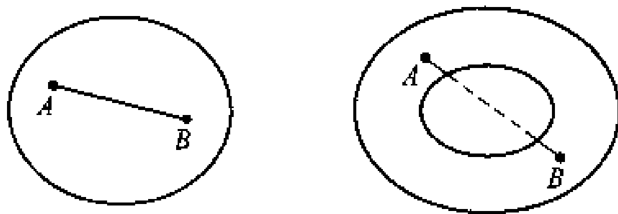


图 2-32 圆为凸集，环为非凸集

如果博弈有 n 个参与者，参与者 i 有 k 个纯策略，那么策略方程 R 的定义域为 $[0,1]^{k\times n}$，它表示$[0,1]$的 $k\times n$ 重笛卡尔积。例如函数 $f(x,y)$，$x,y\in A$，那么函数的定义域就为 $A^{1\times 1}$，即 x,y 都从集合 A 中取值。由于混合策略的取值为$[0,1]$，显然$[0,1]$为非空凸紧集，所以 $[0,1]^{k\times n}$ 也是非空凸紧集。在古诺模型中，它的定义域虽然与上面讲的有所不同，即纯策略有无穷多个，但策略 q（产量）的取值为$[0,a]$，它同样是非空凸紧集。这表明证明纳什均衡真正需要的，并不是"有限的"要求，而是策略空间必须为紧致集合。如果一个参与者有两个纯策略 x 和 y，那么所有混合策略集合（策略空间）就是满足 $x+y=1$，x 和 y 都大于等于零的直线。如果有三个纯策略 x,y,z，那么所有混合策略集合（策略空间）就是一个满足 $x+y+z=1$，x,y,z 都大于等于零的一个三角平面。如图 2-31所示。

（2）最优反应对应是上半连续的。根据前面的介绍，我们已经知道混合策略空间是非空凸紧集，并且 v-M-N 收益函数是混合策略的线性函数，因而它是连续的和凹的（因为偏好为理性），这就保证了根据最大化 v-M-N 收益函数得到的最优反应对应必然是上半连续和非空凸集，这在数学中被称为最大化定理。

综上所述，最优反应方程 $R(p_1,\cdots,p_n)$满足不动点定理的两个条件：策略空间为非空凸紧集，最优反应对应为上半连续并且是非空凸集。因而存在着不动点，而不动点就是纳什均衡。需要指出的是，纳什定理只表明纳什均衡存在的充分条件，而不是必要条件，因而即使不满足以上这两个条件，博弈也可能存在纳什均衡。其次，纳什定理只表明存在纳什均衡，而不排除存在着多个纳什均衡的可能性，而多重均衡会使得我们无法有效地对博弈的结果作出预测。

关键词

完全信息静态博弈　博弈的基本式　严格优策略　严格优策略均衡　优策略　严格劣策略　劣策略　重复剔除严格劣策略法　（纯策略）纳什均衡　严格均衡　非严格均衡　对称博弈和对称均衡　混合策略　v-M-N偏好　理

性的 v－M－N收益函数 （混合策略）纳什均衡 纳什定理 不动点 非空凸紧集 上半连续 不动点定理

内容要点

1. 完全信息静态博弈是指博弈参与者同时行动，并且每个参与者关于收益的信念都是确定的。也可以这样来理解完全信息静态博弈，即所有参与者都不存在信息优势和顺序优势，因而完全信息静态博弈是最简单的博弈模型。

2. 博弈的基本式由三个要素组成：参与者集合、策略空间集合和偏好，简记为 $G=\{S,u\}$。

3. 纳什均衡是指一个策略组合，这个策略组合对于所有参与者的所有策略而言都是最优的，也可以理解为是博弈的一种稳定状态。纳什均衡包含着严格优策略均衡和重复剔除严格劣策略均衡，它是一个较一般的均衡概念。

4. 求解纳什均衡最一般的方法就是写出博弈参与者的最优反应函数（或对应），各参与者最优反应函数的交集即为纳什均衡。无论策略为离散还是连续的博弈，最优反应函数都适用。

5. 利用对称博弈的特性可以简化对博弈的分析。

6. 引入混合策略使得纳什均衡的适用范围更为广阔，而根据纳什定理，只要博弈是有限的，就一定存在（混合或纯）的纳什均衡。

7. v－M－N偏好是理性的，除了要满足非对称性和负传递性，还需要满足替代公理和阿基米得公理，同时 v－M－N收益函数满足线性变换。

进一步阅读

关于 v－N－M偏好非常详细的说明可以参阅 Kreps（1990），纳什定理的证明可以参考 Nash（1950，1951）或者 Fudenberg 和 Tirole（1991），对于不动点定理较全面的介绍可参阅 Mas－Colell，Whinston 和 Green（1995，数学附录）。

第三章

纳什均衡运用举例

学习博弈论最好的方法就是对其加以运用。在这一章里将通过一些经典模型来学习如何分析一个博弈，以及如何求解博弈的纳什均衡解。

第一节　伯川德模型

在前面曾介绍过古诺模型，古诺模型中垄断厂商的策略为产量，如果策略不是产量而是价格又会如何呢？在伯川德（Bertrand，1883）模型中，垄断厂商选择的策略不是产量而是价格，其他方面与古诺模型类似。其内容大致如下：n 个厂商共同生产一种完全同质的产品，$n \geqslant 2$。每个厂商确定价格，然后根据市场需求来生产产品。如果有厂商的价格高于其他厂商，那么其销售额为 0；如果其价格最低，那么他将生产所有的市场份额；如果最低价格不止一家厂商，那么市场份额在它们之间平分。该博弈的基本式 $G = \{S, u\}$ 如下：

（1）参与者集合：n 个寡头厂商，$N = \{1, 2, \cdots, n\}$。

（2）策略空间：策略为价格 p_i，策略空间为 $S_i = [0, p^*]$，p^* 表示一个充分大的价格。实际上假设价格无穷大也无所谓。

（3）偏好和收益函数。寡头厂商的收益函数就是其利润函数，即

$$\pi_i(p_i, p_{-i}) = \begin{cases} 0 & p_i > p_{-i} \\ p_i(D(p_i)/m) - c_i(D(p_i)/m) & p_i \leqslant p_{-i} \end{cases}$$

其中 $D(p_i)$ 表示市场需求函数，m 表示选择最低价格的厂商数，如果只有厂商 i 的价格最低，那么 $m=1$。$p_i > p_{-i}$ 表示 p_i 大于任意一个其他厂商的价格 p_j，其他类似。

我们不妨从简单的开始，设 $n = 2$，$c_i(q_i) = cq_i$，$D(p) = a - p$，当 $p > a$ 时，$D(p) = 0$，$a > c$。由此可得寡头厂商 i 的利润函数为

$$\pi_i(p_i, p_j) = \begin{cases} (p_i - c)(a - p_i) & p_i < p_j \\ \frac{1}{2}(p_i - c)(a - p_i) & p_i = p_j \\ 0 & p_i > p_j \end{cases}$$

那么伯川德模型的纳什均衡是什么呢？

可由如下思路得到寡头厂商 i 的最优反应对应[1]。

如果 $p_j<c$，那么可知当 $p_i\leqslant p_j$ 时，厂商 i 的利润为负数；而 $p_i>p_j$ 时，利润为0；显然利润为0好过利润为负。由此得到最优反应对应为

$$B_i(p_j)=p_i>p_j,\quad p_j<c$$

如果 $p_j=c$，那么可知当 $p_i<p_j$ 时，厂商 i 的利润为负数；而 $p_i\geqslant p_j$ 时，利润为0；显然利润为0要好于利润为负。由此得到最优反应对应为

$$B_i(p_j)=p_i\geqslant p_j,\quad p_j=c$$

如果 $p_j>c$，那么可知 $p_i<p_j$ 时，厂商 i 的利润函数为 $(p_i-c)(a-p_i)$；而 $p_i=p_j$ 时，厂商 i 的利润函数为 $(p_i-c)(a-p_i)/2$；而 $p_i>p_j$ 时，利润为0，显然这时选择 $p_i<p_j$ 是最优策略。由于假设价格是连续变化的，所以 p_i 可以无限接近 p_j，因而从数学意义上而言，最优反应对应可能是不存在的，即最优反应对应为空集，$B_i(p_j)=\varnothing_j$。具体而言，我们需要对 $p_j>c$ 的情况做具体分析。

可知利润函数 $(p_i-c)(a-p_i)$ 为二次型，因而存在一个 p^m 使得利润最大（这里先不考虑寡头厂商 j）。我们也可以对其求一阶条件，得

$$p^m=\frac{c+a}{2}$$

由此可知，当 $p_j>p^m$ 时，厂商 i 的最优策略是 $p_i=p^m$；当 $p^m\geqslant p_j>c$ 时，厂商 i 的最优策略是无穷接近 p_j，即为空集。那么最优反应对应为

$$B_i(p_j)=\begin{cases}p^m & p_j>p^m\\ \varnothing & p^m\geqslant p_j>c\end{cases}$$

总结最优反应对应如下：

$$B_i(p_j)=\begin{cases}p_i>p_j & p_j<c\\ p_i\geqslant p_j & p_j=c\\ \varnothing & p^m\geqslant p_j>c\\ p^m & p_j>p^m\end{cases}$$

可知伯川德模型为对称博弈，因而根据对称博弈的特性，可得寡头厂商 j 的最优反应对应如下：

---

[1] 我们约定俗成地规定，除非特指，否则最优反应对应既指最优反应函数，也指最优反应对应。

$$B_j(p_i)=\begin{cases}p_j>p_i & p_i<c\\ p_j\geqslant p_i & p_i=c\\ \varnothing & p^m\geqslant p_i>c\\ p^m & p_i>p^m\end{cases}$$

两个寡头垄断厂商的最优反应对应如图 3-1 所示。

根据图 3-1，可知伯川德模型存在唯一的纳什均衡 (c,c)，即每个寡头厂商的价格等于单位成本 c，利润为 0，这个结果显著不同于古诺模型，而是等同于完全竞争。

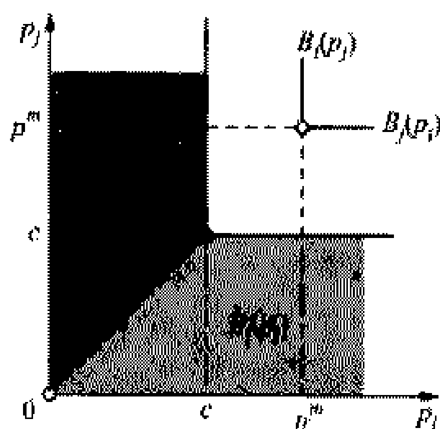


图 3-1 最优反应对应和纳什均衡

现在考虑 $n>2$ 的情况，显然当 n 越来越大时，伯川德的纳什均衡会发生变化，但利润为 0 的特性仍然保持不变，即伯川德模型对于厂商数是中性的。而古诺模型中，均衡产量是 n 的函数，即 $q^*=(a-c)/(n+1)$，当 n 越来越大时，利润越来越接近于 0，这时古诺模型收敛于伯川德模型。

㊣ 习题 3.1 请你证明当 $n>2$ 时，伯川德模型的纳什均衡中至少有两个厂商的价格等于 c。

㊥ 习题 3.2 假设伯川德模型中的两个寡头厂商的单位成本是不同的，不失一般性，$a>c_1>c_2$，那么纳什均衡是什么？如果货币的最小单位为 Δ，纳什均衡又是什么？

㊣ 习题 3.3 设成本函数为 $c_i=f+cq_i$，$c_i(0)=0$，这时寡头厂商 i 的利润函数为 $\pi_i=(p_i-c)(a-p_i)-f$，f 为正但小于 $(p_i-c)(a-p_i)$ 的最大值，$i=1,2$。求该伯川德模型的纳什均衡。

第二节 多党竞选

西方资本主义发达国家在政治制度上基本都实行多党制。从政治主张来看，基本可以划分为保守主张和激进主张，因而一个党派如何选择自己的政治立场，将会关系到能够吸引多少选票，同时党派之间的"立场"选择存在明显的策略依存性。再如，一条街道上两个相互竞争的超市，如何选择位置将会影响吸引到的顾客数量，因为顾客通常更愿意选择离家较近的超市购买东西，因而两个超市如何选址就是一个策略依存的过程。类似的情况还有很多，博弈论中将这类

模型称为豪泰林(Hotelling,1924)模型。为了简单,这里先考虑两党制的情况。

假设保守主张用 0 来表示,激进主张用 1 来表示。选民分布在保守主张和激进主张之间,不妨设 m 为[0,1]中的中位数,即[0,m]区间与[m,1]区间上的选民都为半数。一个简单的假设就是选民服从均匀分布。对于任一选民而言,其偏好是选择越接近自己政治主张的候选人效用越高,即如图 3-2 所示。图 3-2 中,x^* 代表选民的政治主张,越接近 x^*,选民的效用越大。上面假设选民的偏好是对称的,即偏向保守还是偏向激进,其效用减少都一样。实际上,选民是否均匀分布,偏好是否对称,对我们的分析来讲并不重要。

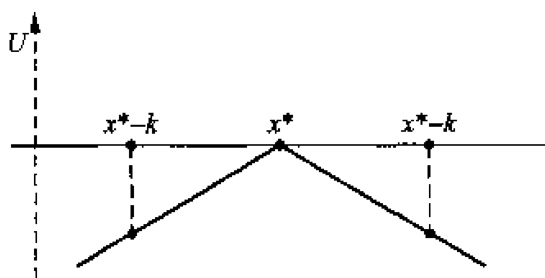


图 3-2 选民的偏好是对称的

在这种情况下,两党就需要确定某种政治主张(实际上就是选择立场)以最大限度地吸引选民投它的票,谁获得的选票多谁就获胜。两党争夺选民的情况可以用图 3-3 来说明。

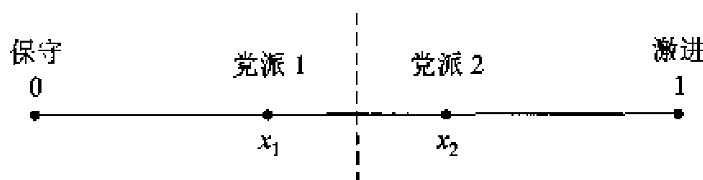


图 3-3 两党争夺选民

如果党派 1 的政治主张(策略)为 x_1,它较偏向于保守主张,所以它能获得保守选民的选票,即得到分布在[0,x_1]的选民的支持。由于假设选民分布在[0,1]的区间,所以它获得的投票数可用 x_1 来表示。对于党派 2 而言,它的政治主张为 x_2,较偏向激进主张,所以它能获得激进选民的选票,即获得分布在[x_2,1]的选民的支持。它所获得的投票数可用 $1-x_2$ 来表示。对于分布在[x_1,x_2]的选民而言,则被两个党派平分,图 3-3 中的虚线的左边为支持党派 1 的选民,虚线的右边为支持党派 2 的选民。显然,给定党派 1 的策略 x_1,党派 2 的策略 x_2 越接近 x_1,获得的选民就会越多;对于党派 1 同样存在类似的情况。那么在这个两党竞选中纳什均衡是什么?该博弈的基本式 $G=\{S,u\}$ 如下:

(1) 参与者:党派 1 定义为 1,党派 2 定义为 2,$N=\{1,2\}$。

(2) 策略空间:党派 1 的策略空间为 $S_1=[0,1]$,策略 $x_1\in[0,1]$;党派 2 的策略空间为 $S_2=[0,1]$,策略为 $x_2\in[0,1]$。

（3）偏好和收益函数：任何一个党派都希望获胜，其次希望能够平分秋色，最差就是竞选落败。由于赢得的选民多即为获胜，因而两个党派的收益实际上是获得的选民数的函数；但选民数又取决于两个党派的政治主张，因而收益函数最终是两党的政治主张的函数。这里不妨设赢的收益为 w，势均力敌为 a，失败为 0。根据已知条件，党派 1 的收益函数 $u_1(x_1,x_2)$ 为

$$u_1(x_1,x_2)=\begin{cases}0 & x_1<x_2\leqslant m \text{ 或 } x_1>x_2\geqslant m\\ a & x_1=x_2=m\\ w & m\geqslant x_1>x_2 \text{ 或 } m\leqslant x_1<x_2\end{cases}$$

根据上述收益函数就能得到相应的最优反应对应。当 $x_2<m$ 时，党派 1 的最优策略 x_1 为大于 x_2，小于 $2m-x_2$ 之间的任意数，因为这时党派 1 获得的选民数超过半数，能够赢得选举。当 $x_2=m$ 时，那么党派 1 的最优策略 x_1 就是等于 x_2，因为这样党派 1 能够获得半数选民的支持，不至于落败。当 $x_2>m$ 时，党派 1 的最优策略 x_1 为大于 $2m-x_2$，小于 x_2 之间的任意数，因为这样能获得超过半数选民的支持，从而赢得选举。最终得到党派 1 的最优反应对应如下：

$$B_1(x_2)=\begin{cases}x_1\in\{x_1:x_2<x_1<2m-x_2\} & x_2<m\\ x_1=x_2 & x_2=m\\ x_1\in\{x_1:2m-x_2<x_1<x_2\} & x_2>m\end{cases}$$

豪泰林模型显然是一个对称模型，因而利用对称性可以很容易地得到党派 2 的收益函数和最优反应对应。党派 2 的最优反应对应如下：

$$B_2(x_1)=\begin{cases}x_2\in\{x_2:x_1<x_2<2m-x_1\} & x_1<m\\ x_2=x_1 & x_1=m\\ x_2\in\{x_2:2m-x_1<x_2<x_1\} & x_1>m\end{cases}$$

其表达的意思与党派 1 的完全一样，这里就不再赘述。显然两个最优反应对应的交集即为豪泰林模型的纳什均衡。图 3－4 形象地反映出了豪泰林模型的纳什均衡，该模型不仅是对称的，而且有对称的纳什均衡 (m,m)。

上面这个两党竞选模型虽然简单，但却深刻地说明了为什么西方发达国家无论什么党派上台，长期奉行的政治路线通常为中间路线，即不过分保守，也不过分激进。像美国、英国这些两党制的国家就是

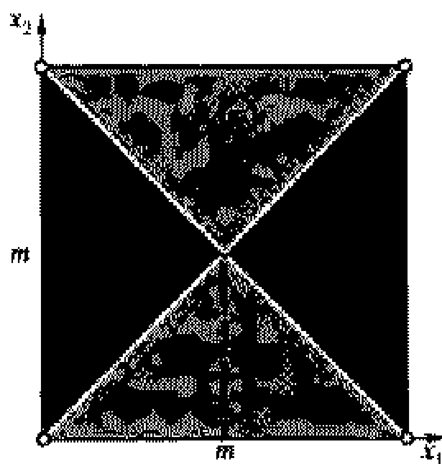


图 3－4　豪泰林模型的最优反应对应与纳什均衡

典型。此外，它还揭示出各党派的政治主张具有趋同性，奉行中间路线的党派最有可能获胜，这意味着真正的政治改革通常难以实行。正像这个模型最早的提出者豪泰林所说的，美国的民主党和共和党总是尽可能地模仿对手，而这种现象在各种电视娱乐节目中更是非常普遍。

两党竞选这个模型还可以进一步变形，例如三党制。在三党制下，纳什均衡不存在，但最为核心的东西并没有变，那就是没有哪一个党派能够获得绝对多数选民的支持，因而通常的政府和议会是以党派联合执政的方式出现，这就解释了为什么实行多党制($i>2$)的国家通常是联合党执政，而很少出现像美国那样的情况。关于三党制的博弈留作习题，由读者进一步分析。

(易) 习题3.4 请建立肯德基和麦当劳的豪泰林模型，并用它来解释为什么肯德基和麦当劳通常建在一起？

(中) 习题3.5 如果竞选的党派超过2个，例如三党制，那么会是什么情况？

(难) 习题3.6 美国的总统选举实行的是代表人团票制度，如果总统候选人在一个州能够获得多数选民的选票，将获得该州的所有代表人团票，获得最多代表人团票的人当选美国总统。这里假设存在 k 个州，用 v 来表示代表人团票，v_1 表示州1的代表人团票，v_2 表示州2的代表人团票，以此类推。并且规定州1的代表人团票数比州2多，州2比州3多，代表人团票最少的就为州 k，即 $v_1>v_2>\cdots>v_k$。总统候选人则选择一个政治主张，即 x_i，各州的选民则根据 x_i 是否离自己的政治主张最近来选择候选人。如果存在两个总统候选人，$k=2$，那么这个博弈的纳什均衡是什么？当 $k=3$ 时，纳什均衡又是什么？

(易) 习题3.7 在垄断竞争中，厂商生产的产品存在着差异性。假设存在两个厂商，他们既可以生产彼此不同质的产品，也可以生产完全一样的产品，即两个厂商生产的产品的差异可大可小，而消费者均匀地分布在差异上。请建立一个豪泰林产品特征模型并分析之。

第三节 事故赔偿法

一旦发生事故，通常需根据法律规定来确认伤害人和受害人之间的责任，以进行相应的赔偿。本节将分析法律规定如何影响伤害人和受害人的行为，以及社会的福利水平。为了便于分析，假设存在两位参与者：伤害人和受害人，显然两位参与者的行为直接关系到事故发生后的损失和需承担的责任，不妨设损失函数为 $L(a_1,a_2)>0$，其中 a_1 和 a_2 分别表示伤害人和受害人的小心程度，显然参与者越小心，那么损失越小，用数学表示即损失函数 L 的偏导数小于0。此外，设 $p(a_1,a_2)$ 为伤害人承担损失的比重，$1-p(a_1,a_2)$ 即为受害人承担损失的

比重。参与者越小心(相对于对手),那么责任就越轻,即 p 的偏导数小于0。函数 L 和函数 p,都是性态良好的。由此就能得到社会、伤害人和受害人的收益函数。

从社会来看当然是希望损失越小越好,因而

$$\max[-a_1-a_2-L(a_1,a_2)] \tag{3-S-1}$$

而伤害人则希望自己的损失越小越好,因而

$$\max[-a_1-p(a_1,a_2)L(a_1,a_2)] \tag{3-S-2}$$

受害人的收益函数则为

$$\max\{-a_2-[1-p(a_1,a_2)]L(a_1,a_2)\} \tag{3-S-3}$$

以上目标函数之所以取负号,原因在于负数的最大值其绝对值最小(损失最小)。

从社会的角度来看,总的损失越小越好,因而对(3-S-1)式求一阶条件得,

$$\frac{\mathrm{d}L}{\mathrm{d}a_1}=\frac{\mathrm{d}L}{\mathrm{d}a_2}=-1$$

根据上式就能得到满足社会最优的 a_1^* 和 a_2^*。

法律对于责任的认定将会影响到参与者各自的行为,通过与 a_1^* 和 a_2^* 的比较就可以对不同的法律规定进行福利分析,而这正是我们所关心的。在美国,长期实行过一种称为**过错加过错**(negligence with contributory negligence)的赔偿原则。为了便于分析,不妨假设当发生事故时,伤害人要么承担全部责任,要么完全不承担责任,即 p 要么等于1,要么等于0。只有当伤害人有过错,而受害人没有过错时,伤害人才承担责任,其他情况下,伤害人都不承担责任。不妨设法律规定的过错标准为 X_1 和 X_2,当 $a_1\geqslant X_1$ 时,伤害人没有过错;当 $a_2\geqslant X_2$ 时,受害人没有过错。因而过错加过错的赔偿原则就可以表示成如下公式

$$p(a_1,a_2)=\begin{cases}0 & a_1\geqslant X_1 \text{ 或 } a_2<X_2\\ 1 & a_1<X_1 \text{ 或 } a_2\geqslant X_2\end{cases}$$

另一个赔偿原则就是所谓的**纯过错**(pure negligence),这时 $X_1>0, a_2=0$,即只有在伤害人没有过错而受害人有完全过错的情况下,伤害人才能免责,否则将承担完全责任。还有一个赔偿原则是绝对**责任**(strict liability),即 X_1 无穷大,而 $a_2=0$,即无论如何伤害人都要承担责任。在 $p=1$ 或0的情况下,a_1^* 和 a_2^* 不仅使社会损失最小,同样使得伤害人和受害人的损失最小。因为在这种情况下,(3-S-1)式与(3-S-2)式仅相差一个 a_2,而与(3-S-3)式仅相差一个 a_1,在求(3-S-1)的偏导数时,a_2 和 a_1 分别被看作常数,因而(3-S-1)的偏导数就分别与(3-S-2)和(3-S-3)的一阶条件相同,即三个收益函数的最优解完全相同。

我们将事故赔偿法概括为一个策略式博弈 $G=\{S,u\}$：

（1）参与者集合：1 代表伤害人，2 代表受害人，$N=\{1,2\}$。

（2）策略空间：策略为 a_i，策略空间为 $S_i=[0,+\infty)$，$i\in N$。

（3）收益函数：根据不同的赔偿原则，将会有不同的收益函数。不妨从最简单的开始。

第一种情景：纯过错。

不妨规定 $X_1=a_1^*$，这时伤害人的收益函数为

$$u_1(a_1,a_2)=\begin{cases}-a_1 & a_1\geqslant a_1^* \text{ 且 } a_2=0\\ -a_1-L(a_1,a_2) & a_1<a_1^* \text{ 或 } a_2>0\end{cases}$$

受害人的收益函数为

$$u_2(a_1,a_2)=\begin{cases}-a_2 & a_1<a_1^* \text{ 或 } a_2>0\\ -L(a_1,a_2) & a_1\geqslant a_1^* \text{ 且 } a_2=0\end{cases}$$

在这种情况下，伤害人的最优反应函数为

$$B_1(a_2)=\begin{cases}a_1^* & a_2=0\\ a_1^{\#} & a_2>0\end{cases}$$

如果受害人小心程度 $a_2=0$，那么伤害人的最优策略就是令 $a_1=a_1^*$；如果受害人的小心程度 $a_2>0$，那么伤害人就需要求使得 $-a_1-L(a_1,a_2)$ 最大的 $a_1^{\#}$。根据已知条件可知，在给定 $a_2<a_2^*$ 的情况下①，$a_1^{\#}>a_1^*$，只有当 $a_2=a_2^*$ 的时候，$a_1^{\#}=a_1^*$。原因在于，当受害人的小心程度不等于 a_1^* 时，伤害人需要付出更多的小心以防止损失增加。由于 (a_1^*,a_2^*) 使得社会总损失最小，所以同样使得伤害人和受害人的损失最小。可以用图 3－5 加以说明。

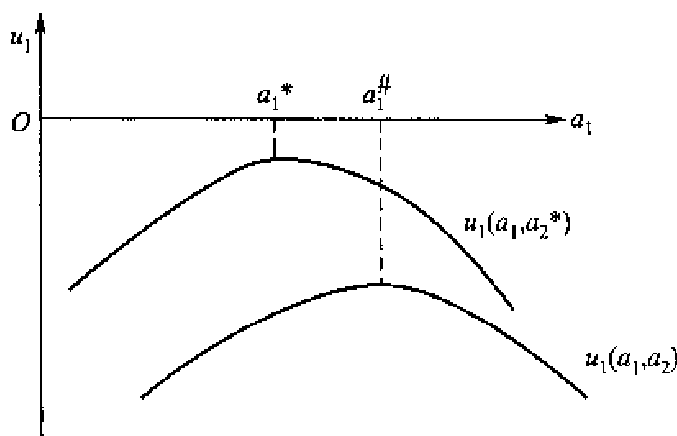


图 3－5　给定 $a_2<a_2^*$，使 u_1 最大的 a_1

① 实际上每个参与者都不会选择大于 a^*，因为那是一个严格劣策略。

对于受害人而言，最优反应函数为

$$B_2(a_1)=\begin{cases}0 & a_1<a_1^*\\ a_2^{\#}>0 & a_1\geqslant a_1^*\end{cases}$$

由于 $a_1<a_1^*$ 时，受害人最优策略为0，而受害人选择0，伤害人的最优策略就为 a_1^*，所以当 $a_1<a_1^*$ 时，不存在均衡。当 $a_1=a_1^*$ 时，受害人的最优策略为 $a_2^{\#}>0$，而受害人选择 $a_2^{\#}$，伤害人的最优策略就为 $a_1=a_1^*$。由于策略是连续的，因而当 $a_1=a_1^*$ 时，受害人可以选择一个无穷接近0，但不等于0的 $a_2^{\#}$，所以该博弈不存在纯纳什均衡解。但如果假设 a 的最小单位为 Δ，那么纳什均衡就为 $(a_1^{\#},\Delta)$。显然，$(a_1^{\#},\Delta)\neq(a_1^*,a_2^*)$，这表明由此形成的均衡结果并不是社会最优的结果。在中国，交通法规中的赔偿原则类似于纯过错，通常伤害人都需要承担一定责任，只有当受害人完全过错，而伤害人完全没有过错时，才可能免责。

接下来分析第二种情景：过错加过错。

从感情上来讲，通常人们很难接受过错加过错这种原则，因为只有在伤害人完全过错，并且受害人完全没有过错的情况下，才能进行赔偿。但出乎意料的是，过错加过错的赔偿原则却会导致社会福利最大。这时伤害人和受害人的收益函数分别为

$$u_1(a_1,a_2)=\begin{cases}-a_1 & a_1\geqslant a_1^*\text{ 或 }a_2<a_2^*\\ -a_1-L(a_1,a_2) & a_1<a_1^*\text{ 且 }a_2\geqslant a_2^*\end{cases}$$

$$u_2(a_1,a_2)=\begin{cases}-a_2 & a_1<a_1^*\text{ 且 }a_2\geqslant a_2^*\\ -a_2-L(a_1,a_2) & a_1\geqslant a_1^*\text{ 或 }a_2<a_2^*\end{cases}$$

给定受害人的策略 $a_2<a_2^*$，那么伤害人的最优策略是选择 $a_1=0$；如果 $a_2\geqslant a_2^*$，那么伤害人的最优策略就是选择 $a_1=a_1^*$。即，最优反应函数为

$$B_1(a_2)=\begin{cases}0 & a_2<a_2^*\\ a_1^* & a_2\geqslant a_2^*\end{cases}$$

给定伤害人的策略 $a_1<a_1^*$，那么受害人的最优策略就是选择 $a_2=a_2^*$；如果伤害人的策略为 $a_1=a_1^*$（伤害人不可能选择 $a_1>a_1^*$ 的值，因为它是严格劣策略），那么受害人的最优策略仍为 $a_2=a_2^*$。因而，受害人的最优反应函数为

$$B_2(a_1)=\begin{cases}a_2^* & a_1<a_1^*\\ a_2^* & a_1=a_1^*\end{cases}$$

因而存在唯一的纳什均衡 (a_1^*,a_2^*)，而这恰恰就是社会福利最大时所要求

的小心度。(a_1^*,a_2^*)不仅是社会最优解,而且还是伤害人和受害人的最优解,这就解释了为什么在西方国家交通秩序非常好,原因就在于过错加过错的赔偿原则;而中国的交通秩序通常要差,原因就在于几乎无论什么情况下,行人和骑车人都能获得一定赔偿。需要强调的是,这里的分析假设伤害人和受害人完全平等并高度抽象,否则乱用这种制度安排就有可能导致特权和非常不公的社会结果,因而它不具任何政策和法律上的建议性,仅是一种分析而已。

(易) 习题 3.8 在过错加过错原则下,如果 $X_1 < a_1^*$,那么纳什均衡会改变吗?

(易) 习题 3.9 请你分析绝对责任情景下的博弈。

第四节 公共地悲剧

18 世纪初期,资本主义生产方式在英国得到了确立,并开始迅速发展。而它的最初动力是纺织品市场需求的日益扩大,使得毛纺工业迅速发展,纺织机、蒸汽机相继得到运用,直接导致第一次工业革命。羊毛作为原料,市场需求巨大。但当时英国农村的土地属于公社所有,牧场更是如此。在这种情况下,土地公有实际上成为有效生产羊毛的一个重要障碍,最终导致马克思称之为的“羊吃人”的圈地运动。1739 年,休谟(Hume)在其所写的一篇文章中指出,公共牧场实际上会导致过度放牧,从而为英国圈地运动的产生提供了一种说明。在经济学中,公共地悲剧现在已成为说明私有制存在理由的经典模型之一。

“公共地悲剧”讲述的是一个村社有 $n \geq 2$ 个牧民。村社的公共牧场向每一个村社的成员开放。在春天,牧民同时选择所需要饲养的羊群数。假设羊群数是连续可分的。每年夏季,所有的牧民都在村社的公共牧地上放牧自己的羊群。如果第 i 个牧民拥有的羊群数为 g_i,那么村社所有的羊群数就为 $G = g_1 + \cdots + g_n$。买和饲养一只羊的成本为 c,这里假设它是一个常数。每只羊的价值为羊的总头数的函数,即 $v(G)$。由于每只羊都需要一定的绿地才能生存,因而村社的牧场所能供养的羊群数就有一个上限,定义为 G_{max},并假设当所饲养的羊群数超过牧场所能供养的上限时,每只羊的价值为零(即羊群过多造成牧场退化,由于没有足够的饲料,从而引起羊群死亡),即当 $G < G_{max}$ 时,$v(G) > 0$;当 $G \geq G_{max}$ 时,$v(G) = 0$。由于一开始饲养的羊群数很少,因而新增一只羊对已经饲养的羊来说不会造成太大的损失,但这种损失是递增的,即当 $G < G_{max}$ 时,$v'(G) < 0$,$v''(G) < 0$①。一阶导数为负表明随着羊的数量增加,羊的价值会下降;二阶导数为负则表明随着羊的数量增加,羊的价值下降是递增的。如图 3-6 所示。

---

① v'表示一阶导数,v''表示二阶导数。

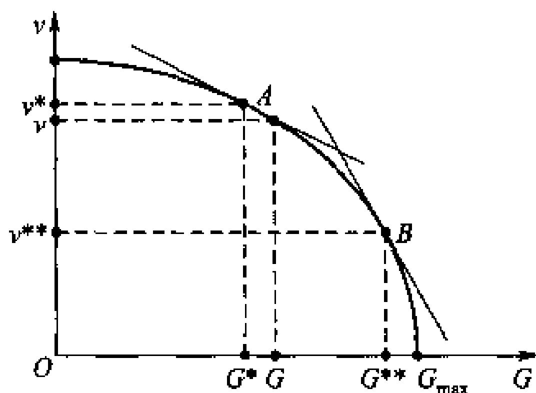


图 3-6 羊的价值以递增的速度下降

图 3-6 中的曲线代表每只羊的价值，它随着 G 的增加而下降，因而它的斜率为负，即 $v'(G)<0$。同时，在 A 点的斜率的绝对值要小于在 B 点的斜率的绝对值，这表明随着 G 的增加，羊的价值的下降幅度越来越大，即 $v''(G)<0$。例如，$G^*G=G^{**}G_{max}$，当 G^* 增加到 G 时，羊的价值从 v^* 下降到 v，下降幅度（即边际量）为 v^*-v，而当 G^{**} 增加到 G_{max} 时，羊的价值从 v^{**} 下降到 0，下降幅度为 v^{**}。在羊的增量相同的情况下，价值下降的边际量却不一样。现在将该博弈转化为基本式 $G=\{S,u\}$：

（1）参与者集合：定义牧民为 $1,2,\cdots,n$；$N=\{1,2,\cdots,n\}$。

（2）策略空间：牧民 i 的策略为 g_i，策略空间 $S_i=[0,G_{max}]$，$i\in N$。

（3）收益函数：当其他的牧民所放牧的羊群数为 $(g_1,\cdots,g_{i-1},g_{i+1},\cdots,g_n)$ 时，牧民 i 从放牧 g_i 的羊群所获得的收益为

$$u_i(g_1,\cdots,g_i,\cdots,g_n)=g_iv(g_1+\cdots+g_{i-1}+g_i+g_{i+1}+\cdots+g_n)-cg_i \tag{3-G-1}$$

每个牧民追求收益最大化，即

$$\max_{g_i}\quad u_i(g_1,\cdots,g_i,\cdots,g_n)$$

假如 $(g_1^*,\cdots,g_n^*)$ 是纳什均衡的话，对于每一个牧民 i 而言，在给定其他人的策略组合 $(g_1^*,\cdots,g_{i-1}^*,g_{i+1}^*,\cdots,g_n^*)$ 下，g_i^* 一定会使(3-G-1)式最大化。

最大化的必要条件（一阶条件）为

$$v(g_i+g_{-i}^*)+g_iv'(g_i+g_{-i}^*)-c=0 \tag{3-G-2}$$

最大化的充分条件（二阶条件）为

$$2v'(g_i+g_{-i}^*)+g_iv''(g_i+g_i^*)<0$$

其中 $g_{-i}^*=g_1^*+\cdots+g_{i-1}^*+g_{i+1}^*+\cdots+g_n^*$。利用对称性，可知(3-G-2)式是每一个牧民的收益最大化的充要条件。根据纳什均衡的定义，g_i^* 一定是(3-G-2)式的解。将 g_i^* 代入(3-G-2)式，并将所有牧民的必要条件相加，得

$$nv(G^*)+G^*v'(G^*)-nc=0 \qquad (3-G-3)$$

用 n 除(3 - G - 3)式,得

$$v(G^*)+\frac{1}{n}G^*v'(G^*)-c=0 \qquad (3-G-4)$$

其中 $G^*=g_1^*+\cdots+g_n^*$。(3 - G - 4)式是纳什均衡下,所有牧民所养羊之和必须满足的条件。

与私人利益最大化相对立的是社会的收益最大化,社会的目标是使下式最大化

$$\max_{0\leqslant G<\infty} Gv(G)-Gc$$

一阶条件为

$$v(G^{**})+G^{**}v'(G^{**})-c=0 \qquad (3-G-5)$$

二阶条件为

$$2v'(G^{**})+G^{**}v''(G^{**})<0$$

G^{**} 就是社会的最优解。现在我们需要证明的是 $G^*>G^{**}$,即公共地存在过度利用,怀海特(Whitehead,1948)把这种情况称为“公共地悲剧”。

证明用反证法。假设 $G^*\leqslant G^{**}$。

因为 $v'<0$

所以 $v(G^*)\geqslant v(G^{**})$ (第1项)

又因为 $v''<0$

所以 $0>v'(G^*)\geqslant v'(G^{**})$

所以 $|v'(G^*)|\leqslant|v'(G^{**})|$ (第2项)

根据假设 $G^*\leqslant G^{**}$

所以 $\frac{G^*}{n}<G^{**}$ (第3项)

根据(第1项)、(第2项)、(第3项)比较(3 - G - 4)式和(3 - G - 5)式,可得出(3 - G - 4)式大于(3 - G - 5)式的结论。这显然与(3 - G - 4)和(3 - G - 5)都等于零相矛盾,所以 $G^*>G^{**}$。命题得证。

$G^*>G^{**}$ 表明:① 公共牧场被过分利用,资源没有达到有效配置;② 拥有羊群越多的牧民损失越多,因为每只羊的价值 $v(G)$ 是 G 的减函数。因而从效率的角度来看,圈地运动必然要求实行公共牧场的私有化,并按照羊头数的多少来划分土地专属权的大小。历史的发展恰恰与上述的理论分析是一致的,由于羊头数少的牧民占绝大多数,因而就出现了被马克思称之为“羊吃人”的现象。然而,如果没有圈地运动这场“农业革命”,随后的英国工业革命是不可能发生的。

“公共地悲剧”反映出了极其深刻的哲理——大家都拥有(公地)的自由,却

导致了大家都受损的结局。这也正是哲学家怀海特把其称作为"公共地悲剧"的原因，在怀海特的著作中这样写道："悲剧的实质不在于不幸，而在于（这种不幸是由大家共有）这一神圣性（所造成的）。"[①]要避免公共地悲剧的发生，如果排除私有化的可能性，靠技术是不可能解决的，它实际上取决于人们崇高的思想境界。

㊁ 习题 3.10　假设一条河流的上游存在一个造纸厂，河的下游存在一个渔场，把其概括为一个博弈并分析之。

第五节　争议仲裁

在现实的经济活动中，经常会发生各种争议，例如劳资争议、合同争议、消费争议等。一旦争议发生，就需要社会的仲裁机构（通常为民事法庭）来加以仲裁。仲裁的主要形式有两种：协议仲裁和最后仲裁。根据民法原则，争议双方应首先自愿协商解决，只有争议双方无法达成协议，仲裁机构才进行最后仲裁。在中国城市化进程的过程中，一个突出的矛盾就是房屋拆迁中的赔偿金问题，这里我们引人一个赔偿金最后仲裁模型，实际上这个模型可以推广到各式各样的争议仲裁博弈中。

假定参与争议的双方一为政府，另一为居民户，争议因补偿金而起。首先，政府和居民户同时开出自己希望的赔偿金水平，分别用 w_g 和 w_f 表示。如果 $w_g > w_f$，显然不存在争议。如果 $w_g < w_f$，双方就存在补偿金争议，并报仲裁机构进行仲裁。最后，仲裁机构会在两者之中选择其一作为结果。这个博弈实际上属于后面要讲到的动态博弈，但由于仲裁机构超脱于该博弈，在这个博弈中不存在收益，所以可以把该博弈看作是静态博弈来处理，因为政府和居民户报价之后，结果实际上已经决定了（不存在走后门、拉关系这样的可能）。在博弈中，把类似于仲裁机构的这种参与者称为自然，它参与博弈但却没有自己的利益。假定仲裁机构本身对补偿金有自己认为合理的值，用 x 来表示。在政府和居民户报价 w_g 和 w_f 之后，仲裁机构只是简单选择距 x 最为接近的要价：如果 $x < (w_g + w_f)/2$，仲裁者将选择 w_g；如果 $x > (w_g + w_f)/2$ 则选择 w_f，如果 $x = (w_g + w_f)/2$，选择哪一个无关紧要，因为政府和居民户报价的期望值相等。仲裁机构的仲裁原则如图3－7所示。仲裁机构遵循的这个仲裁原则是公共信息。

① Whitehead A. N. 1948. "Science and the Modern World." New York: Mentor: 17

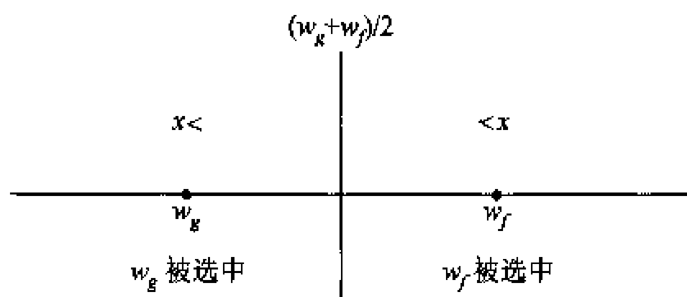


图 3－7 可能的仲裁结果

对于 x，只有仲裁机构知道，政府和居民户不知道。但 x 服从一定的概率分布是公共信息，其累积分布函数为 $F(x)$，概率密度函数为 $f(x)$。根据仲裁原则，如果双方的要价分别为 w_g 和 w_f，那么双方推断 w_g 被选中的概率为 $Pr\{w_g\}$ 和 w_f 被选中的概率为 $Pr\{w_f\}$，用数学表示出来为：

$$Pr\{w_g\} = Pr\left\{x < \frac{w_g + w_f}{2}\right\} = F\left(\frac{w_g + w_f}{2}\right)$$

和

$$Pr\{w_f\} = 1 - F\left(\frac{w_g + w_f}{2}\right)$$

据此，期望的补偿金水平为

$$w_g Pr\{w_g\} + w_f Pr\{w_f\} = w_g F\left(\frac{w_g + w_f}{2}\right) + w_f\left[1 - F\left(\frac{w_g + w_f}{2}\right)\right]$$

这里需要说明的是，$(w_g >)w_f^{\#}$ 相对于 $(w_g =)w_f$ 是居民户的严格劣策略；从政府来说，$w_g^{\#}(>w_f)$ 相对于 $w_g(=w_f)$ 是政府的严格劣策略，因而从理论上而言不会出现 $w_g > w_f$ 的情况。

政府的目标是获得期望补偿金水平最小化的仲裁结果，居民户则设法使其最大化。政府和居民户之间形成了一种博弈关系，读者现在应当能够写下该博弈的基本式，这里我们不再重复。

根据纳什均衡的定义，在给定居民户策略 w_f 的条件下，政府追求支付最小化，即

$$\min_{w_g} w_g F\left(\frac{w_g + w_f^*}{2}\right) + w_f^*\left[1 - F\left(\frac{w_g + w_f^*}{2}\right)\right]$$

一阶条件为

$$\frac{1}{2}(w_f^* - w_g) f\left(\frac{w_g + w_f^*}{2}\right) = F\left(\frac{w_g + w_f^*}{2}\right) \qquad (3-Z-1)$$

在给定政府策略 w_g^* 的条件下，居民户则追求收益的最大化，即

$$\max_{w_f} w_g^* F\left(\frac{w_g^* + w_f}{2}\right) + w_f\left[1 - F\left(\frac{w_g^* + w_f}{2}\right)\right]$$

一阶条件为

$$\frac{1}{2}(w_g^* - w_f)f\left(\frac{w_g^* + w_f}{2}\right) = \left[1 - F\left(\frac{w_g^* + w_f}{2}\right)\right] \qquad (3-Z-2)$$

如果纳什均衡为(w_g^*, w_f^*)，它一定同时满足(3－Z－1)和(3－Z－2)式。这意味着(3－Z－1)式和(3－Z－2)式的左边项相等，因而有

$$F\left(\frac{w_f^* + w_u^*}{2}\right) = \frac{1}{2} \qquad (3-Z-3)$$

这意味着政府和居民户报价的平均值等于仲裁机构 x 分布的均值，这说明政府和居民户都努力想与仲裁机构的偏好吻合。因为如果偏离 x 越远，那么对手的报价就越有可能被选中。将(3－Z－3)代入(3－Z－1)或(3－Z－2)，得

$$w_f^* - w_g^* = \frac{1}{f\left(\frac{w_g^* + w_f^*}{2}\right)} \qquad (3-Z-4)$$

政府和居民户报价之差等于仲裁机构 x 分布的均值点的概率密度函数的倒数。为了更好地理解这一结果，假设 x 服从均值为 m，方差为 σ^2 的正态分布(也可以假设 x 服从均匀分布等)，那么密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{\frac{(x-m)^2}{2\sigma^2}\right\} \qquad (3-Z-5)$$

根据(3－Z－3)可得

$$\frac{w_g^* + w_f^*}{2} = m \qquad (3-Z-6)$$

将式(3－Z－6)代入(3－Z－4)，并根据(3－Z－5)式得

$$w_f^* - w_g^* = \frac{1}{f(m)} = \sqrt{2\pi}\sigma \qquad (3-Z-7)$$

根据式(3－Z－6)和(3－Z－7)就能得到纳什均衡的要价为

$$\left(w_g^* = m - \frac{\sqrt{2\pi}\sigma}{2},\ w_f^* = m + \frac{\sqrt{2\pi}\sigma}{2}\right)$$

这表明双方的均衡要价以仲裁机构 x 的期望值 m 为中心对称，且要价之差随双方对仲裁机构 x 的不确定性(即 σ^2)的提高而增大。对这一均衡结果的直观解释非常简单，一个更为“不合理”的要价(即居民户更高的要价或政府更低的出价)一旦被仲裁机构选中就会给自己带来更高的收益，但其被选中的可能性却会相应降低，因而博弈的双方都需进行权衡。当政府和居民户认为仲裁机构 x 的不确定程度增加(即 σ^2 变大)时，双方的要价就会更“不合理”，因为一个更“不合理”的要价与仲裁机构的 x 有较大差别的可能性变小了。相反，如果不确定性较小，双方都不会开出离 x 太远的报价，因为仲裁机构选择离 m 最近的

报价的可能性非常大。如果不存在不确定性,双方则都会开出 x 的报价。

争议仲裁这个博弈令人吃惊的地方在于:房屋拆迁的补偿金实际上与居民户在拆迁过程中的损失无关,而是完全取决于仲裁机关的偏好。如果仲裁机关与政府的偏好一致,那么实际上唯一的均衡就是 w_g,这就从一定程度上解释了中国近年来房屋拆迁过程中为什么矛盾如此突出。

㊂ 习题 3.11　试比较分析,仲裁机关的 x 是根据房产的市场价格来确定与 x 由政府规定这两种情况会导致怎样不同的结果?存不存在寻租的可能?

第六节　合作性谈判

这里我们考虑一个博弈模型:有一元钱,两兄弟分,两兄弟必须同时给出各自认为应得的份额。不妨假设兄弟1给出的份额为 x_1,而兄弟2给出的份额为 x_2。博弈的规则是,如果两兄弟给出的份额为 $x_1+x_2>1$,那么两兄弟只能得到0;如果两兄弟给出的份额之和 $x_1+x_2\leqslant 1$,那么他们各自得到自己提出的份额。这实际上是一个简单的合作性博弈,合作成功对双方都有利,而合作失败则双方都受损。根据前而对纳什均衡的定义,很容易知道这个合作性博弈存在无穷多个纳什均衡。凡是 $x_1,x_2\geqslant 0$,并且 $x_1+x_2=1$ 的所有 (x_1,x_2) 策略组合都是纳什均衡。因为如果给定兄弟1的策略 x_1,那么兄弟2的最优策略为 $1-x_1$。因为如果 $x_2>1-x_1$,那么双方都只能得到0,显然这比得到 $1-x_1$ 要差;如果 $x_2<1-x_1$,显然也比得到 $1-x_1$ 要差,因此只有 $x_2=1-x_1$ 才是兄弟2的最优策略。对于兄弟1而言,给定兄弟2的策略 x_2,兄弟1的最优策略为 $x_1=1-x_2$,分析同理,不再赘述。但在现实生活中,经常观察到的一个理象是,人们在分配利益时往往出现的均衡是平均分配,即兄弟1和兄弟2平均分配,各得1/2元钱。为什么会如此?这就引出了纳什最早开创的合作性谈判研究。

合作性谈判的模型,是由纳什独立发展出来的,并由此开启了合作性博弈(cooperative games)研究的大门。他把一个谈判问题极端地抽象化,除去所有的谈判细节和过程,而专注于讨论谈判的空间和谈判的结果之间的关系。具体讲,一个双人谈判问题,是两个人在所有可能的谈判结果所产生的效用空间里,去协商选出一个效用组合来。如果谈判不成,那么他们将各自得到一个谈判破裂的效用。纳什问的问题是,在不同的效用可能及协商破裂的可能下,怎样去选出一个公平的谈判结果。

定义 3.1　一个双人谈判问题(two - person bargaining problem),是一个 R^2 里的集合 X,以及 X 中的一个点 d,记作 (X,d)。其中 X 代表所有谈判可能的结果下,谈判者的所有可能效用的集合。d 是谈判破裂点(disagreement point),代

表当他们谈判破裂时，各自所得的效用。令 Σ 是所有谈判问题的集合。也就是说，$\Sigma=\{(X,d)\mid X\subset R^2,d\in X\}$。

假设 X 是凸集合(convex set)和紧集(compact set)。再假设存在 $x\in X$ 使得 $x_i>d_i,i=1,2$。即假设 X 中存在让两个人效用都大于谈判破裂时的效用，这样谈判才有意义。

定义 3.2 一个谈判问题的解 F 是一个从 Σ 映射到 R^2 的函数。也就是说，对于任意一个 $(X,d)\in\Sigma,F(X,d)$ 是 X 中的一个点。它代表的意思是，面对 (X,d) 这样一个谈判问题时，谈判者从最后的分配所得到的效用组合是 $F(X,d)$。

从定义 3.2 可以看出，纳什所找的，不是一个单一的分配值，面是一个一般的、普遍的分配原则。在这个原则下，不论遇到什么分配问题，都可以用这个原则去分配。同时，纳什认为一个"好"的分配原则(即谈判的解)，应该满足下列四个公设：

A1 效率性(efficiency)

假如有一点 $x>F(X,d)$，则 $x\notin X$。这就是说谈判的解 F 必须是帕累托效率(Pareto efficiency)的分配。

A2 尺度不变性(invariance)

所谓不变性是指谈判的解与效用采用什么度量单位无关。也就是说，对于任何的 $u=(u_1,u_2)\in R^2$，令 $f_i(u_i)=a_iu_i+b_i$，且 $a_i>0$，其中 $i=1,2$。假若 u_i 代表第 i 个人的效用值，那么 $f_i(u)$ 等于是把第 i 个人的效用做一个线性变换。当我们把一个人的效用做线性变换时，他的效用其实并没有改变，只是用以衡量的单位变了。因此他们的谈判解的唯一改变，也应该只是完全相同的单位改变而已。[①] 令 $f=(f_1,f_2)$。尺度不变性即是要求 $F(f(X),f(d))=f(F(X,d))$。这说明如果把原来问题的所有效用值，都用 f 去变换，那么变换之后的问题的解，应该也是原先问题的解的相同变换。

A3 无关选择的独立性(independence of irrelevant alternative)

假设 $Y\subset X$，且 $F(X,d)\in Y$，则 $F(Y,d)=F(X,d)$(见图 3-8)。[②] 这个公设的意思是说，假如 $F(X,d)$ 是较大谈判空间 X 下的最优解，那么 $F(X,d)$ 也一定是较小空间 Y 的最优解。这表明 X/Y 的部分其实是无关紧要的。因此，当谈判被限制在较小的 Y 里面选择时，原先的解 $F(X,d)$ 也是现在的解。

A4 对称性(symmetry)

如果原先的问题是完全对称的，那么两个人所得到的效用应该完全相等。

---

① 微观经济学中偏好是序数，而不是基数。

② 这个公设事实上并不是百分之百合理和毫无争议的。

也就是说，若$(x_1,x_2)\in X\Rightarrow(x_2,x_1)\in X$；且$d_1=d_2$，那么$F_1(X,d)=F_2(X,d)$（见图3-9）。对称性实际上是说，如果两人完全平等（金钱、地位等），那么谈判的结果对于两人而言必然是一样的。谈判结果不一样只可能要么有人是傻瓜，要么双方的实力不对等。实力不对等则可以用$d_1\neq d_2$或是不对称的X表示。

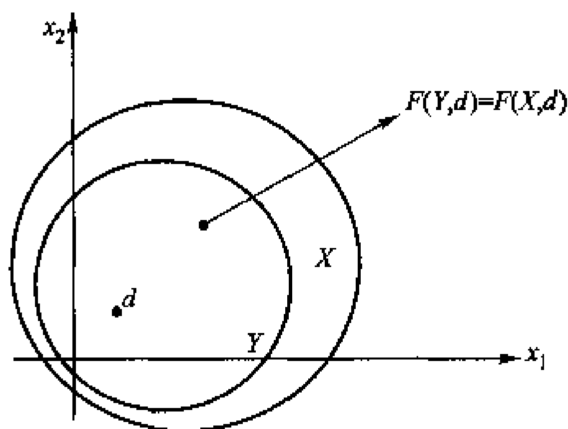


图3-8　无关选择的独立性

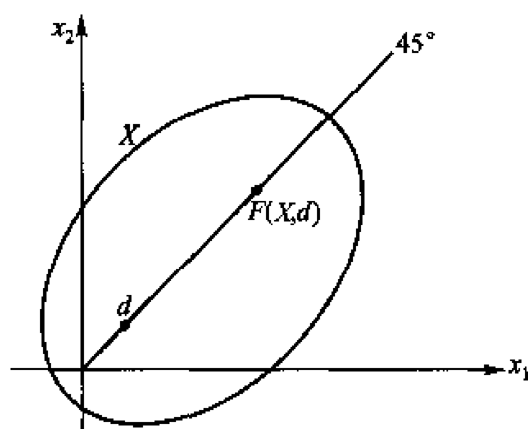


图3-9　对称性

定理3.1（Nash）　上述双人谈判问题中只有一个解满足公设**A1—A4**，那就是

$$F(X,d)=\arg\max_{(x_1,x_2)\in X}(x_1-d_1)(x_2-d_2) \qquad (a)^{①}$$

即$F(X,d)$其实是以(d_1,d_2)为中心的双曲线与X的切点（见图3-10）。

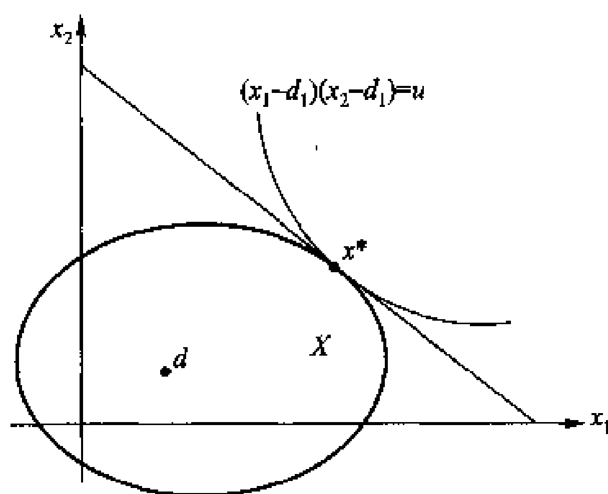


图3-10　合作性谈判的解

证明：我们分两步来证明。

首先证明如果$F(X,d)=\arg\max_{(x_1,x_2)\in X}(x_1-d_1)(x_2-d_2)$，则$F$一定满足公设**A1—A4**。

① arg max在数学中表示最优解的集合。

(1) 已知 $F(X,d)$ 是上述问题的最大化解。现在不妨假设存在一个 $y>F(X,d)$，并且 $y\in X$，那么可知 $F(X,d)$ 不是上述问题的最大化解，而这与已知条件 $F(X,d)$ 是上述问题的最大化解相矛盾。

(2) 设 $f=ax+b, a>0$。那么 $x^*\in F(f(X),f(d))=\arg\max(f_1(x_1)-f_1(d_1))(f_2(x_2)-f_2(d_2))=\arg\max a^2(F(X,d))$，同时 $x^*\in f(F(X,d))=\arg\max f[(x_1-d_1)(x_2-d_2)]=\arg\max aF(X,d)+b$。即最优分配方案 x^* 不随效用单位进行线性变换而改变。

(3) 不妨设 $x^*\in F(X,d)$，已知 $Y\subset X$，并且 $x^*\in Y$，那么 x^* 一定是集合 Y 中的解，如果不是那就意味着存在一个 $y^*\in F(Y,d)$，有 $y^*>x^*$，而这意味着 x^* 不会是上述问题的最大化解($y^*\in X$)。显然与已知条件矛盾。

(4) 显然 $d_1=d_2$，且 X 对称，$\max(x_1-d)(x_2-d)\rightarrow x_1^*=x_2^*$。如果不是，除非 d_1 不等于 d_2，或是 X 不是对称的。

较困难的部分，是证明若有任何一个解 F 满足公设 **A1—A4**，那么 F 一定满足(a)式。令 $x^*=\arg\max_{(x_1,x_2)\in X}(x_1-d_1)(x_2-d_2)$。换言之，$x^*$ 是 X 与 (d_1,d_2) 为中心的双曲线的切点。我们要证明如果任一个解 F 满足公设 **A1—A4**，那么 $F(X,d)=x^*$。

先把问题标准化。也就是说，我们要找一个变换，把 x^* 的坐标转换成 $(1,1)$，d 的坐标转换成 $(0,0)$。令

$$g_i(z)=\frac{z-d_i}{x_i^*-d_i},\quad i=1,2$$

则很容易得出 $g(x^*)=(1,1)$，且 $g(d)=(0,0)$。不但如此，原先和 X 相切于 x^* 点上的切线，在经过变换后会成为一条和横轴成 45°，且和 X 相切于 $g(x^*)$ 点的线(见图 3-11)。由于 X 是紧致集，因此 $g(X)$ 必然也是。同时，可以把 $g(x^*)$ 这一点上的切线延长，划出一对称于 45°线的等腰三角形 $\triangle ABC$，使得 $g(X)\subset\triangle ABC$。我们先把 F 用在 $(\triangle ABC,g(d))$ 这个谈判问题上。由于 $\triangle ABC$ 以 45°线对称，且 $g_1(d)=g_2(d)=0$，因此由公设 **A4** 可知 $F(\triangle ABC,g(d))$ 一定在 45°线上。再由公设 **A1** 可知 $F(\triangle ABC,g(d))$ 是 45°线上唯一满足帕累托效率的点。因此，$F(\triangle ABC,g(d))=(1,1)=g(x^*)$。又由于 $g(X)\subset\triangle ABC$，因而 $F(\triangle ABC,g(d))=g(x^*)\in g(X)$，因此由公设 **A3** 可知 $F(g(X),g(d))=F(\triangle ABC,g(d))=(1,1)$。再令 $f_i=g_i^{-1}$，也就是让 f_i 为 g_i 的反函数，则 $f_i(y)=(x_i^*-d_i)y+d_i$，很明显 f_i 是 y 的线性转换。因此由公设 **A2** 可知 $F(X,d)=F[f(g(X)),f(g(d))]=f(F(g(X),g(d))=f(1,1)=x^*$，上式的第二个等式来自公设 **A2**。因此证明了 $F(X,d)$ 一定等于 x^*，而 x^* 正是满足(a)式的解。

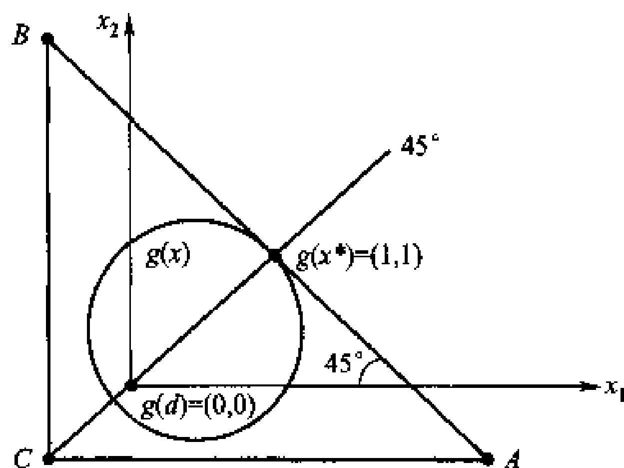


图 3-11 谈判空间的标准化

根据定理 3.1 可知，分钱博弈虽然有无穷多个纳什均衡，但只有唯一一个纳什均衡是公平的分配，即两兄弟各得 1/2 的钱。令 $d=(0,0)$，即如果两兄弟没有合作成功，那么他们得到 0 的收益，那么 $\{x_1=1/2, x_2=1/2\}$ 是(a)式唯一的最大化解。

定理 3.1 是一个非常优美、对称，且合乎情理与直觉的定理。它从逻辑的意义上证明了什么能称为公平(没有对错之分)，揭示了为什么人们会认为平均分配就是公平。纳什对谈判问题的处理和证明方式，都深深地影响并笼罩着后来合作博弈的发展。例如他使用的公理化(axiomitization)，一直到今天，仍然是合作博弈研究者所使用的标准方法。

㊟ 习题 3.12 什么是公平？公设不同显然公平的定义就不同，如何理解公平与效率？平均分配真的与效率是矛盾的吗？无法实现平均分配就在于人们之间的地位不平等吗？

第七节 报案

我们经常能在新闻中看到这样的报道，一个犯罪嫌疑人在众目睽睽之下从事犯罪活动，而周围却有很多人袖手旁观。又如，一位儿童落水，同样有很多人无动于衷。面对这些社会现象，有各种各样的解释，这里仅从博弈的角度来探讨其原因。

设有 n 个人观察到一起犯罪活动。每个人都希望能够通知警察，制止犯罪活动，但是他们都偏好别人打电话(因为打电话报警可能要付出成本，例如，需要付电话费，需要报身份证号等)。如果没有人打电话，那么自己打，总要好过受良心的谴责。具体而言，该博弈的基本式如下：

(1) 参与者集合：n 个旁观者，$N=\{1,\cdots,n\}$。

(2) 策略空间：$S_i=\{C,D\}$，C 代表打电话，D 代表不打电话，$i=N$。

(3) 偏好：对于第 i 个参与者，如果别人打电话而自己不打电话，收益为 v，如果自己打电话那么收益为 $v-c>0$，如果没有人打电话，那么收益为 0。

根据纯纳什均衡的定义，容易得到{一个人打电话，而其他人不打电话}是纳什均衡，没有人打电话不是纳什均衡，这显然与现实中观察到的现象不完全吻合。那么没有人打电话会不会也是一个纳什均衡呢？如果引人混合策略，那么答案是肯定的。

不妨假设每个人打电话的概率都为 $1>p>0$，那么 $1-p$ 就是不打电话的概率，因而混合策略为 $(p,1-p)$。根据前面所学知识，我们知道如果参与者运用混合策略，那么概率为正的纯策略下的期望值都应该相等。如果第 i 个旁观者打电话，那么他的收益为 $v-c$；如果他不打电话，那么可能出现两种情况：一种是没有人打电话，另一种是至少有一个其他人打电话，因而其期望值为

$$0\times Pr\{\text{没有其他人打电话}\}+v\times Pr\{\text{至少有一个其他人打电话}\}$$

在均衡条件下必有

$$v-c=0\times Pr\{\text{没有其他人打电话}\}+v\times Pr\{\text{至少有一个其他人打电话}\}$$

没有人打电话的概率实际上等于 1 减去至少有一个其他人打电话的概率，即

$$Pr\{\text{没有其他人打电话}\}=1-Pr\{\text{至少有一个其他人打电话}\}$$

将其带人上式整理后就得

$$v-c=v\times(1-Pr\{\text{没有其他人打电话}\})$$

或

$$c/v=Pr\{\text{没有其他人打电话}\}$$

根据已知条件，$Pr\{\text{没有其他人打电话}\}=(1-p)^{n-1}$，由此得

$$p=1-(c/v)^{1/(n-1)}。$$

根据命题 2.7 和命题 2.8，立知该博弈除了存在 n 个纯纳什均衡之外，还存在唯一一个混合纳什均衡，即每一个旁观者都以 $p=1-(c/v)^{1/(n-1)}$ 的概率打电话，以 $(c/v)^{1/(n-1)}$ 的概率选择不打电话。报案这个博弈有意思的地方在于：在混合纳什均衡中，确实存在没有人打电话的情况，而且没有人打电话的概率随着旁观者人数的增加面增加。这一点通过下列计算就能看出来：

$$\begin{aligned}Pr\{\text{没有人打电话}\}&=Pr\{\text{参与者 } i \text{ 没有打电话}\}\times Pr\{\text{其他人没有打电话}\}\\&=(1-p)^n=(c/v)^{n/(n-1)}\end{aligned}$$

对上式求关于 n 的一阶导数，得

$$\frac{\mathrm{d}Pr}{\mathrm{d}n}=\ln(c/v)\cdot(c/v)^{n/(n-1)}\cdot\left(\frac{-1}{(n-1)^2}\right)>0$$

即随着旁观者的增加，没有人打电话的概率是增加的，由于至少有一个人打

电话的概率等于1减去没有人打电话的概率,因而旁观者越多越有可能出现见死不救的情况。为什么会出现这种情况?实际上,如果只存在一个旁观者,那么打电话的概率为1,但在存在多个旁观者的情况下,任意一个旁观者不打电话的概率为$(c/v)^{1/(n-1)}$。由于c/v小于1,当n增加时,不打电话的概率$(c/v)^{1/(n-1)}$会变大,原因就在于旁观者之间产生了相互依赖的心理,总是寄希望于别人打电话报警。虽然每个人都性本善$(v-c>0)$,但又存在一点点私心$(v>v-c)$,最终就有可能出现最不幸的结局。

社会冷漠与见义勇为

在1964年3月,美国纽约市发生了一起凶杀案,被害人被歹徒残酷杀害,令人感到震惊的是38位旁观者在长达半个小时的时间里,面对被害人的呼救无动于衷,甚至无一人报警。该案的发生震动了美国社会,引起了整个美国社会的大讨论和反思,报案模型就是来自于这一真实案件。2007年中央电视台的3·15晚会中报道了湖南省的一个城市的河堤边的同一个地方先后有20多位儿童落水身亡,面对悲剧的一再重复同样表现出了令人震惊的冷漠,面对一位失去女儿的母亲的奔走呼告,无人采取相应的回应,使得悲剧一而再,再而三的发生。对于这些社会现象,社会学、伦理学、心理学等学科会给出各自的解释,但在博弈论看来,社会冷漠的产生根源于人们的自利心,尽管人是好人,但也是追求自利的人。如果我们崇尚见义勇为的精神,那么就不应忘记制度的重要,在制度上保证奖励见义勇为者和处罚冷漠者,要比单纯地依靠所谓的高尚精神来的牢靠,如果见义勇为者总是吃亏者,那么见义勇为的行为就会在社会上彻底消失。

(易) 习题3.13　请用建立模型来解释为什么很多人对落水儿童见死不救。(假设水很浅但却足以使儿童死亡)如果要让人们见义勇为应该如何做?

第八节　专家诊断

当人们遇到问题时,通常会征求专家的意见以便解决问题。面对专家的诊断,你可能接受也可能拒绝,原因在于专家有动机会把小问题说成是大问题以便

获得较高的收益。我们不妨将这种情景模型化。假设一个消费者在消费某种产品时,产品无法正常工作,这个问题有可能是小问题,也可能是大问题,出现小问题的概率为 $1>r>0$。专家知道这个问题是大问题还是小问题(不存在误诊),但他可以讲真话,也可以讲假话(把小问题说成是大问题)以获得更大利益。消费者不能确定是大问题还是小问题,但可以拒绝专家的建议以寻求进一步的解决办法。大问题大修,小问题小修,小修不能解决大问题,但大修却能解决大问题和小问题。显然,消费者对于专家小修的建议始终采纳,因为小问题大修能使专家获得更多利益(但专家却没有这样做,说明专家是诚实的),而小修并不能解决大问题(消费者将不会支付任何维修费用)。小问题小修和大问题大修都将使专家获得正常利润 π,但如果小问题大修将使专家获得 $\pi'(>\pi)$ 的利润。如果大修,那么消费者支出 H,如果小修支出 $L<H$。如果消费者拒绝专家大修的建议,那么在专家说真话的情况下,需要支付更高的代价 $H'>H$;在专家说假话的情况下,需要支付 L',其中 $H>L'>L$。在上述这些假设下,专家的行动就为诚实和欺骗,而消费者的行动则为接受和拒绝,其博弈的基本式如图 3-12 所示。

		消费者	
		接受(q)	拒绝($1-q$)
专家	诚实(p)	$\pi, -rL-(1-r)H$	$r\pi, -rL-(1-r)H'$
	欺骗($1-p$)	$r\pi'+(1-r)\pi, -H$	$0, -rL'-(1-r)H'$

图 3-12　专家诊断

(诚实,接受):这时无论大问题还是小问题,专家的收益都为 π,对于消费者而言,小问题出现的概率为 r,所以其期望支付为 $-rL-(1-r)H$。

(诚实,拒绝):这时专家大修的建议会被拒绝,所以其期望收益为 $r\pi+(1-r)\times 0$,而消费者的期望支付为 $-rL-(1-r)H'$。

(欺骗,接受):这时专家始终建议大修,而消费者始终接受专家的建议,因而专家的期望收益为 $r\pi'+(1-r)\pi$,而消费者支付 H。

(欺骗,拒绝):这时专家始终建议大修,而消费者始终拒绝,所以专家的收益为 0,而消费者的支付为 $-rL'-(1-r)H'$。

根据纳什定理可知,该博弈一定存在纳什均衡。不妨设专家诚实的概率为 p,消费者接受的概率为 q。给定消费者的混合策略 q,专家概率为正的纯策略的期望收益应该都相等,即

$$q\pi+(1-q)r\pi = q[r\pi'+(1-r)\pi]$$

$$q = \frac{\pi}{\pi'} \qquad (3-zh-1)$$

即当 $q=\pi/\pi'$ 时，专家的混合策略为 $1 \geq p \geq 0$。

当 $q<\pi/\pi'$ 时，专家的混合策略为 $p=1$，即始终选择诚实。

当 $q>\pi/\pi'$ 时，专家的混合策略为 $p=0$，即始终选择欺骗。

反之，给定专家的混合策略 p，那么消费者概率为正的纯策略的期望收益同样应该都相等，即

$$-p[rL+(1-r)H]-(1-p)H$$
$$=-p[rL+(1-r)H']-(1-p)[rL'+(1-r)H']$$

整理后得

$$p = \frac{H-rL'-(1-r)H'}{r(H-L')} \qquad (3-zh-2)$$

从上式可以看出，当 $H<rL'+(1-r)H'$ 时，$p<0$，显然这是不可能的。进一步的分析显示，当 $H<rL'+(1-r)H'$，表明消费者如果拒绝，那么实际支出要大于接受专家的建议（尽管专家欺骗），因而消费者只能始终接受专家建议，由于消费者始终接受专家的建议，因而专家的最优策略是始终欺骗，即 $p=0$。

当 $H>rL'+(1-r)H'$ 时，表明消费者拒绝专家的建议寻求进一步的帮助其期望支出要小于始终接受，因而专家的最优策略只能是令 $p>0$，即采取一种混合策略，有时讲真话，有时讲假话，以便消费者摸不清楚。

另外，式(3-zh-2)又可写成 $p=1-\frac{(1-r)(H'-H)}{r(H-L')}$，即 $p<1$。

总结上述分析可知，当 $H<rL'+(1-r)H'$ 时，存在唯一一个的纯策略纳什均衡（欺骗，接受）。当 $H>rL'+(1-r)H'$ 时，则存在唯一一个混合策略纳什均衡（$q=\frac{\pi}{\pi'}$，$p=\frac{H-rL'-(1-r)H'}{r(H-L')}$）。当 $H=rL'+(1-r)H'$ 时，在专家欺骗的情况下，消费者接受或拒绝其收益都一样，因而纳什均衡也为（欺骗，接受）。

为了更为形象，我们利用最优反应函数来进一步验证上述分析。

专家的最优反应函数为

$$B_1(q) = p = \begin{cases} 1 & q < \pi/\pi' \\ [0,1] & q = \pi/\pi' \\ 0 & q > \pi/\pi' \end{cases}$$

而消费者的最优反应函数则需要考虑两种情况。当 $H<rL'+(1-r)H'$ 时，消费者的最优反应为 $B_2(p)=q=1, p\in[0,1]$，如图 3-13 所示。

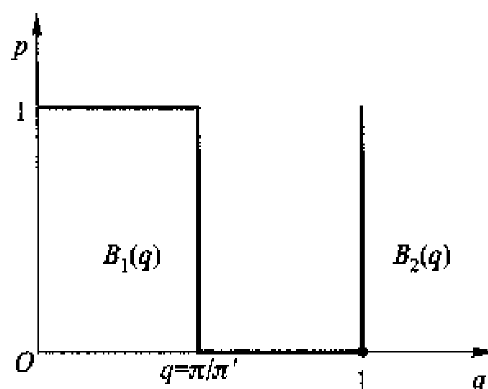


图 3-13 $H<rL'+(1-r)H'$时的纳什均衡

当$H>rL'+(1-r)H'$时，消费者的最优反应函数为

$$B_2(p)=q=\begin{cases}0 & p<[H-rL'-(1-r)H']/r(H-L') \\ [0,1] & p=[H-rL'-(1-r)H']/r(H-L') \\ 1 & p>[H-rL'-(1-r)H']/r(H-L')\end{cases}$$

如图 3-14 所示。纳什均衡为(欺骗，接受)和$\left(p=\dfrac{H-rL'-(1-r)H'}{r(H-L')},q=\dfrac{\pi}{\pi'}\right)$。关于参数如何影响均衡，读者可以通过对 p 和 q 求一阶导数来加以分析，这里不再重复。有意思的是消费者的 q，它取决于专家的正常利润和欺骗时的利润之比，如果两者的比越大，说明正常利润和欺骗利润非常接近(表明专家越不值得欺骗)，因而消费者接受专家建议的可能性越大。而专家的 p 则其决于 H 与 $rL'+(1-r)H'$的接近程度，两者越接近表明消费者越无所谓，因而专家欺骗的动机越强烈。

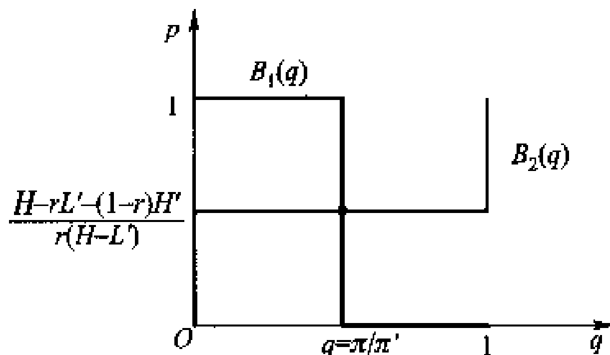


图 3-14 当$H>rL'+(1-r)H'$时的纳什均衡

(易) 习题 3.14　在专家诊断模型中，如果$H=rL'+(1-r)H'$时，那么它的混合纳什均衡是什么？

(易) 习题 3.15　在专家诊断模型中，专家不会欺骗，但却会出现误诊，将小问题诊断成大问题，大修能够解决大问题和小问题。请你建模分析之并找出其纳什均衡，这个模型的均衡与原始专家模型的均衡存在区别吗？

第九节　连续行动下的混合策略纳什均衡

一个有趣的问题是，如果行动为连续时，存不存在混合策略及其纳什均衡。为了回答这个问题，我们不妨从一个简单的例子开始。

【例 3.1】　全付拍卖(all－pay auction)

设两个人竞拍 100 元人民币，他们的报价为 0 到 100 元之间的整数，单位为分。如果报价为 $b_1>b_2$，那么参与者 1 得到 $10\ 000-b_1$ 的收益，而参与者 2 得到 $-b_2$ 的收益；如果两人报价一样，即 $b_1=b_2$，那么各得 $(5\ 000-b_i)$ 的收益($i=1,2$)。由于人民币最小的货币单位为分，因而存在着 10 001 种报价。显然，该博弈不存在纯纳什均衡。为了证明这一点，我们不妨先设两人的报价相等，即 $b_1=b_2$，那么这是否为纳什均衡？答案是否定的。因为如果 $b_1<10\ 000$，那么只要参与者 2 报 $b_2=b_1+1$，其收益将从 $5\ 000-b_2$ 变为 $10\ 000-b_2$；如果 $b_1=10\ 000$，那么参与者 2 报 $b_2=0$，其收益将从 $-5\ 000$ 变为 0。所以，$b_1=b_2$ 不可能是纳什均衡。其次，如果 $b_1\neq b_2$，不失一般性，设 $b_1>b_2$，那么参与者 2 只要提高报价就可以使其收益由 $-b_2$ 变为 $5\ 000-b_2$ ($b_1=10\ 000$) 或是 $10\ 000-b_2$ ($b_1<10\ 000$)。所以，$b_1\neq b_2$ 也不可能是纳什均衡。综上所述，全付拍卖不存在纯纳什均衡。但根据纳什定理，可知该博弈是有限博弈，至少存在一个混合策略纳什均衡。现在就来求这个混合策略纳什均衡。由于该博弈存在 10 001 种纯策略(以分为单位)，所以不妨先设第二个参与者按照 1/10 001 的概率取 0 到 100 元人民币之间的任意整数，那么根据命题 2.8 可知，如果第 1 个参与者也使用混合策略的话，那么其每一个概率为正的纯策略的期望值都应该相等，即当 $b_1\in\{0,1,\cdots,10\ 000\}$ 时

$$Pr\{b_2<b_1\}\times(10\ 000-b_1)+Pr\{b_2=b_1\}\times(5\ 000-b_1)+Pr\{b_2>b_1\}\times(-b_1)=c$$

其中 c 为常数。显然，$Pr\{b_2<b_1\}=b_1/10\ 001$，$Pr\{b_2=b_1\}=1/10\ 001$，而 $Pr\{b_2>b_1\}=1-(b_1+1)/10\ 001$，我们可以近似地把 $Pr\{b_2=b_1\}$ 看作 0，这时 $c=0$，整理上式，得

$$\frac{(10\ 000-b_1)b_1}{10\ 001}-\frac{(10\ 000-b_1)b_1}{10\ 001}=0$$

这就证明了我们的猜测，当参与者 2 取 1 到 10 000 的报价的概率都为 1/10 001时，如果参与者 1 也使用混合策略，那么参与者 1 概率为正的纯策略期望收益都相等。

反之，如果参与者 1 取 1 到 10 001 的报价的概率也为 1/10 000，同时参与者

2 也使用混合策略，那么他的概率为正的纯策略期望收益同样相等。

总而言之，全付拍卖存在一个对称混合策略纳什均衡——每个参与者以 1/10 001的概率取 1 到 10 001 之间的整数。实际上，根据命题 2.7，立即可知每个参与者的每个报价的概率为 1/10 001 的混合策略为对称混合策略纳什均衡。

实际上，全付拍卖博弈中策略的单位为分（分为 0.01 元，近似为零），因而我们可以近似地将其看作是连续的，即每个参与者的报价为[0,100]之间的任意实数。显然闭区间[0,100]是紧致的，因而同样存在混合策略纳什均衡。那么如何寻找连续行动下的混合纳什均衡呢？首先需要对命题 2.8 进行扩展，以适应连续行动的要求。

命题 3.1（命题 2.8 在连续策略下的扩展） 在博弈 $G=\{S,u\}$ 中，我们称混合策略组合 p^* 是一个**混合策略纳什均衡**，当且仅当，对于每一个参与者 i 而言，

(a) 如果策略组合(s_i,p^*_{-i})为参与者 i 带来的期望收益小于 p^* 的收益，那么 p_i^* 分配给策略集合 s_i 的概率为 0；

(b) 没有任何一个策略 s_i，能够使得(s_i,p^*_{-i})为参与者 i 带来的期望收益大于 p^* 的期望收益。

命题 3.1 和命题 2.8 的区别就在于：在连续随机变量下，一个纯策略的概率始终为 0，因而所谓混合策略是在行动集合（即区间）的情况下来讨论，因为连续随机变量下只有在区间内的概率才不为 0。例如，在一个简单的两人博弈中，参与者的策略集合即为区间$[x_i,y_i]$，$i=1,2$，混合策略则为区间$[x_i,y_i]$上的积累分布函数 $F_i(s_i)$。如果(F_1,F_2)存在且为混合策略纳什均衡，那么根据命题 3.1，它一定满足下列条件：

(1) 如果参与者 i 的混合策略为 F_i，那么在区间$[x_i,y_i]$之外的概率为 0，即当 $z<x_i$ 时，$F_i(z)=0$；当 $z>y_i$ 时，$F_i(z)=1$。

(2) 当策略为 s_i，而其他参与者的混合策略为 F_j 时，其期望收益

$$\begin{cases} = c_i & x_i \leqslant s_i \leqslant y_i \\ \leqslant c_i & s_i < x_i \text{ 或 } s_i > y_i \end{cases}$$

其中 c_i 为常数。

为了便于理解，我们仍然用全付拍卖来说明，只不过行动为连续的。

【例 3.2】 连续行动下的全付拍卖

两个人竞拍 100 元人民币，他们的报价为[0,100]中的任意实数。如果报价为 $b_1>b_2$，那么参与者 1 得到 $100-b_1$ 的收益，而参与者 2 得到 $-b_2$ 的收益；如果两人报价一样，即 $b_1=b_2$，那么各得$(50-b_i)$的收益$(i=1,2)$。显然，该博弈不存在纯纳什均衡，其证明与例 3.1 中的证明完全一样。

由于行动是连续的，因而需要引入累积分布函数 $F(x)$ 来描述混合策略，即

$$F(x) = Pr\{y < x\} = Pr\{y \leqslant x\}$$

上式的第二个等式在于连续随机变量 y 的 $Pr\{y = x\} = 0$。显然，对于累积分布函数，在前面内容中已经不止一次地用到过，只不过没有说明而已。例如，在例 3.1 中，$Pr\{b_2 < b_1\} = \sum_{b_2=0}^{b_1-1} \frac{b_2}{10\ 001} = \frac{b_1}{10\ 001} = F(b_1)$。现在，不妨设参与者 2 在策略区间 $[a,b]$ 上的混合策略为 $F_2(x)$，那么如参与者 1 也在区间 $[a,b]$ 上采用混合策略的话，其在策略区间 $[a,b]$ 内的任意报价的期望值应该都相等，不妨令其等于常数 c，即

$$F_2(b_1)(100 - b_1) - (1 - F_2(b_1))b_1 = c$$

令 $a = 0, b = 100$，那么可知当报价为 0 时，其期望值始终为 0，所以 c 又等于 0，这就意味着 $F_2(b_1) = b_1/100$，即累积分布函数为均匀分布。由于该博弈是对称博弈，因而存在对称的混合策略纳什均衡，即每一个参与者在均衡下的混合策略为均匀分布。在离散的情况下，均匀分布就意味着每个报价的概率为 1/10 001，即连续下的分析实际上与离散下的分析完全一样。

【例 3.3】 混合策略伯川德模型

我们已经知道伯川德模型存在一个纯纳什均衡，即垄断厂商的价格都等于成本，因而利润为 0，但这与现实并不完全吻合，那么是否存在混合策略纳什均衡呢？假设利润函数为 $(p - c)D(p)$，其关于 p 是递增的，$p \geqslant 0$。实际上 $p < c$ 对垄断厂商来说是严格劣策略，因而垄断厂商不可能采用小于成本的定价。我们将证明，对于每一个 $p_L > c$，在区间 $[p_L, \infty)$ 上都存在一个混合策略纳什均衡，其中每个厂商都使用相同的混合策略 F，$F(p_L) = 0$，当 $p > p_L$ 时，$F(p) > 0$。

根据命题 3.1 可知，对于区间 $[p_L, \infty)$，垄断厂商 1 的期望收益为

$$F_2(p_1) \times 0 + [1 - F_2(p_1)] \times (p_1 - c)D(p_1) = c$$

当 $p_1 = p_L$ 时，其期望收益为 $(p_L - c)D(p_L)$，即 $c = (p_L - c)D(p_L)$。这就意味着

$$F_2(p_1) = 1 - \frac{(p_L - c)D(p_L)}{(p_1 - c)D(p_1)} > 0$$

由于 p_L 为区间的下限，所以 $F_2(p_L) = 0$；又由于利润函数关于 p_1 是递增函数，所以 F_2 关于 p_1 同样是递增的，当 $p_1 \to \infty$ 时，$F_2 \to 1$，所以 F_2 是累积分布函数。又由于该博弈是一个对称博弈，垄断厂商 2 拥有同样的混合策略 F，这就使我们的断言得到了证明。

这里分析的混合策略均衡实际上都是非常简单的博弈，有时候连续行动下的混合策略纳什均衡可能非常复杂而很难分析。

㊁ 习题 3.16　重新考虑连续行动下的全付拍卖,唯一的变动为参与者人数 $n>2$,请你找出一个对称的混合策略纳什均衡。

关键词

豪泰林模型　事故赔偿法模型　过错加过错　纯过错　绝对责任　公共地悲剧　争议仲裁模型　合作性谈判　效率性　尺度不变性　无关选择的独立性　对称性　公理化　报案模型　专家诊断　连续行动下的混合策略

内容要点

1. 本章主要介绍了八个模型,通过这八个模型使读者进一步理解纳什均衡及如何求解纳什均衡,同时使读者领会博弈论是运用什么样的方法和手段来解释各种社会经济现象的。

2. 求解纳什均衡的方法:一是先猜后证;二是利用最优反应函数;三是利用混合策略下各行动的概率(大于0或是等于0)找出所有纳什均衡。这一章主要介绍了利用最优反应函数求解纳什均衡的方法。

3. 本章还给出了求解连续行动下的混合策略的一种简便方法,即运用命题3.1。

4. 命题3.1实际上是离散行动下混合策略特性的一个推广,本质上与命题2.8相同,区别只在于从累积分布上来理解混合策略。

进一步阅读

合作性谈判(bargaining)最早也是由Nash(1950b)提出的,并从此开启了合作性博弈的大门。关于连续策略下的混合策略的讨论可参阅Osborne(2004,第3章)。本章所举例子都可在其他博弈论教科书中找到,例如Gibbons(1992),Osborne(2004)和张维迎(1996)。

第四章

完全信息动态博弈

下象棋是一个典型的完全信息动态博弈，它与石头、剪刀、布这个游戏显然不同。在石头、剪刀、布这个游戏中，参与者都是"同时行动"的，而在象棋这个游戏中，参与者却是"非同时行动"，有先有后，轮流出招。如果我们将石头、剪刀、布这个游戏变成轮流出招，那么后出招的参与者将始终获胜，石头、剪刀、布这个游戏将没有任何意义。反过来，如果我们要求象棋的参与者同时出招，那么象棋这个游戏将无法进行，象棋也不再是象棋。这说明，博弈参与者是否"同时行动"对于博弈而言是至关重要的，改变行动顺序将会改变整个博弈的性质。这就引出了两个问题：一是适用于静态博弈的分析方法能否同样适用于动态博弈；二是如果不适用，那么新的方法又是什么？

对于动态博弈而言，静态博弈中的纳什均衡概念要求过松，完全照搬将会产生一些极不合理的均衡解，原因在于动态博弈具有静态博弈所不具备的一些特性，分析动态博弈可能面临一些新的困难。例如，在静态博弈中，不存在过去也不存在将来，而在动态博弈中，参与者的将来决策，必须取决于现在所作的决策，而现在的决策又必须考虑将来可能出现的情况，即人们常说的"人无远虑，必有近忧"。在这种情况下，策略空间会变得非常巨大，给求解带来很大困难。另一个困难就是所谓的**动态一致性**（dynamic consistency）。所谓动态一致性是指策略不仅在博弈的局部是最优的，而且在整个博弈中也是最优的。现实生活中经常遇到这样的情况，人们现在认为最优的策略，可能在未来发现不过是错误的策略。而动态一致性要求我们在计算动态博弈时，绝对不能犯这种错误。为此，在完全信息动态博弈中引入了一个新的均衡概念——子博弈完美均衡（subgame perfect equilibrium）和一种求解方法——逆向归纳法（backward induction）。实际上，静态博弈只不过是退化的动态博弈，子博弈完美均衡本身就是纳什均衡。

第一节 完美信息与完全但不完美信息

完全信息动态博弈可以分为两类，即完美信息与完全但不完美信息。所谓完美信息博弈，是指博弈中的后行动者始终能够观察到先行动者的行动，因而动态博弈中不存在参与者同时行动这样的情况。而完全但不完美信息博弈，则指动态博弈中，至少存在两个参与者同时行动的情况，因而“后行动者”无法观察到“先行动者”的行动。不妨用两个例子来加以说明。

【例 4.1】 动态囚徒困境

动态囚徒困境与静态版本唯一不同的地方是，囚徒 1 首先行动（招供或沉默），然后被囚徒 2 所观察到，囚徒 2 再相应地采取行动（招供或沉默），随后博弈结束。显然，囚徒 2 作为后行动者始终能够观察到囚徒 1 的行为，因而这是一个最简单的完美信息动态博弈，如图 4－1 所示。

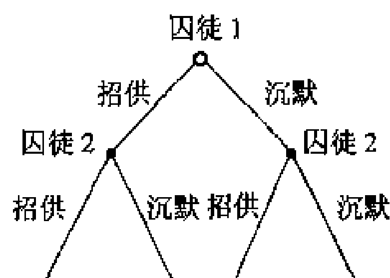


图 4－1 动态囚徒困境

【例 4.2】 取消管制

政府对某个行业进行严格的管制，不准私营企业进行经营。现在政府决定是否取消管制，如若维持管制，那么博弈就此结束；如若取消，那么两个私营企业将同时决定是否进入该行业，然后博弈结束。政府追求社会福利最大，而两个私营企业追求利润最大。该博弈如图 4－2 所示。我们用 1 和 2 来分别代表两个私营企业，进表示进入该行业，退则表示不进入该行业，两个企业的策略组合总共有 4 个。显然，这个博弈中的企业 1 和企业 2 能够观察到政府的行为，但彼此不能观察到各自的行为，因而它是一个完全但不完美的动态博弈。

取消管制博弈也可以用图 4－3 来表示。其中表示企业 2 的两个点被虚线连接起来以表明企业 2 没有观察到企业 1 的行动，由于企业 1 在企业 2 之前，所以企业 1 同样不能观察到企业 2 的行动。但这种表示方法不唯一，也可以把企业 1 和企业 2 调换，表达的意思仍然一样。

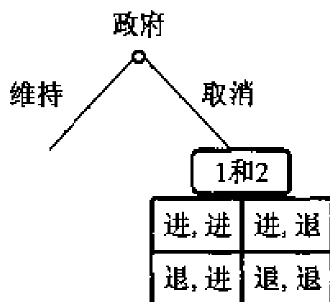


图 4－2 取消管制(1)

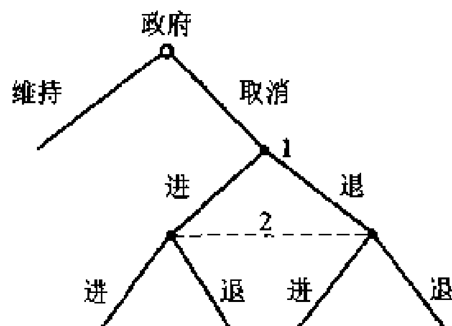


图 4－3 取消管制(2)

定义 4.1　完美信息动态博弈就是不存在同时行动的完全信息动态博弈。

从定义 4.1 可知，完美信息博弈一定是完全信息博弈，但完全信息博弈不一定是完美信息博弈。完美信息博弈比完全信息博弈要求要严格，它不仅要求信息是完全的，而且要求能够观察到对手的行动。因而，完美信息动态博弈是最简单的一种"游戏"。

显然，运用策略式来描述动态博弈会非常不便，特别是当信息不完全时更是如此，为了更简便地描述动态博弈，我们将引入一种新的博弈表达式——扩展式。

第二节　动态博弈的扩展式

如果要完整地描述一个完全信息动态博弈，除了参与者集合、参与者偏好之外，还需要能完整地描绘出博弈的所有过程（即行动序列），以及博弈进行到某个阶段时，哪一个参与者将采取行动。我们把博弈中所有从开始到结束的行动序列称为**全历史**（Terminal history），而用**参与者函数**来表示在每一个全历史上，在博弈进行到某个阶段时谁来行动。因而要完整地描述一个动态博弈，必须具备四个要素：

(1) 参与者集合；

(2) 全历史集合；

(3) 参与者函数；

(4) 偏好。

如果我们把全历史表示成一个行动序列$(a^1, a^2, \cdots, a^K)$（K 为自然数，当 $K \to \infty$ 时，就表示无穷动态博弈），那么$(a^1, a^2, \cdots, a^m)$，其中 $m \leqslant K$，就称为全历史$(a^1, a^2, \cdots, a^K)$的**子历史**（subhistory）。当 $m < K$ 时，$(a^1, a^2, \cdots, a^m)$就是全历史$(a^1, a^2, \cdots, a^K)$的**真子历史**（proper subhistory）。显然，在博弈开始前的历史是一个**空历史**（empty history），因而空历史是所有全历史的真子历史。今后我们将用 $\varnothing$ 来表示空历史，用 h 来表示子历史，而用 H 来表示全历史的集合，而 P 则表示参与者函数。例如，在例 4.1 动态囚徒困境中，存在着四个全历史，即（招供，招供），（招供，沉默），（沉默，招供），（沉默，沉默）。全历史（招供，招供）有三个子历史：空历史 $\varnothing$，（招供），（招供，招供），其中空历史 $\varnothing$ 和（招供）为真子历史。$P(\varnothing)$ = 囚徒 1，表示博弈开始；P(招供) = P(沉默) = 囚徒 2，表示在历史（招供）和（沉默）之后轮到囚徒 2 行动。而在例 4.2 取消管制中，则存在五个全历史，即（维持），（取消，[进，进]），（取消，[进，退]），（取消，[退，进]），（取消，[退，退]）。全历史（取消，[进，进]）同样存在三个子历史：空历史，（取

消),(取消,[进,进])。$P(\varnothing)$ = 政府,P(取消) = {1,2}。其中[进,进]为向量,分别代表企业 1 和企业 2 同时发生的行动①。

有了上面的说明,现在就可以给出完全信息动态博弈的扩展式。

定义 4.2　完全信息动态博弈的扩展式为 $\Gamma = \{N,H,P,u\}$,其中 N 为参与者集合,H 为博弈的全历史集合,即 $H = \{(a^1,a^2,\cdots,a^K)\}$,其中 K 为博弈从开始到结束依次发生的行动次数,行动序列中的每一个 a 都为向量。P 为参与者函数,即 $P(h) = \{i:i \in N\}$, $h \in H$。u 为收益函数,表示博弈参与者的偏好。

与博弈的基本式相比,扩展式没有直接给出博弈参与者的行动集合,原因在于扩展式已经隐含地定义了各参与者在行动时有些什么样的行动可供选择,根据全历史和参与者函数,能很容易地得到各参与者的行动集合。在历史 h 之后,参与者 $P(h)$ 所有可能的行动集合定义为

$A_{P(h)}(h) = \{a_{P(h)}$:(h,a)是一个子历史,$a_{P(h)}$是行动向量 a 的第 $P(h)$ 个元素$\}$

例如,在取消管制博弈中,根据全历史集合 H 和参与者函数 $P(\varnothing)$ = 政府,P(取消) = {1,2},可知 $A(\varnothing)$ = {维持,取消},即政府有两个行动——维持和取消;A_1(取消) = {进,退}和 A_2(取消) = {进,退},即两个企业各有两个行动——进和退。

需要注意的是,在完美信息下,扩展式有三个地方与完全但不完美信息不同。首先,历史 h 由行动向量序列变为行动序列,例如,在取消管制中,历史(取消,[进,进])是一个向量序列,因为企业 1 和企业 2 是同时行动的,如果改成企业 2 后行动,那么就变成(取消,进,进),也就是由一个向量序列便成了单值序列,意思也完全不一样了。其次,参与者函数 $P(h)$ 都是单点映射,对应着唯一一位参与者。最后,就是行动集合 A 可以省略下标,因为 $A(h) = \{a$:(h,a)是一个子历史$\}$。

现在将例 4.1 和例 4.2 的扩展式表达如下:

例 4.1 动态囚徒困境的扩展式为 $\Gamma = \{N,H,P,u\}$,其中:

(1) 参与者集合:囚徒 1 和囚徒 2,$N = \{1,2\}$。

(2) 全历史集合:招供为 C,沉默为 S,H = {(C,C),(C,S),(S,C),(S,S)}。

(3) 参与者函数:$P(\varnothing) = 1$,$P(C) = P(S) = 2$。

(4) 偏好:对于囚徒 1 而言,最好的历史是(C,S),其次为(C,C),然后为(S,S),最倒霉的历史为(S,C)。对囚徒 2 而言,最好的历史是(S,C),其次为(C,C),第三为(S,S),最差为(C,S)。

① 我们约定俗成,参与者 1 的行动排在第一位,参与者 2 的行动排在第二位,以此类推。

例 4.2 取消管制的扩展式为 $\Gamma = \{N,H,P,u\}$，其中：

（1）参与者集合：政府、企业 1 和企业 2，$N = \{1,2,3\}$。

（2）全历史集合：维持为 C，取消为 D，进入为 E，退出为 Q，那么全历史集合 $H = \{(C),(D,[E,E]),(D,[E,Q]),(D,[Q,E]),(D,[Q,Q])\}$。

（3）参与者函数：$P(\varnothing) = 1, P(D) = \{2,3\}$。

（4）偏好：对于政府而言，根据五个历史对应的社会福利进行排序，对于企业 1 和企业 2 而言，则为五个历史对应的利润排序。

【例 4.3】 进入博弈

在一个垄断行业，已经存在一个垄断企业，将其称为在位者，现在有一个新的企业决定是否进入该行业，将其称为挑战者。挑战者首先行动，决定进入（E）或是退出（Q），如果挑战者进入，那么在位者将决定是与挑战者和平共处（A）或是战斗（F）。其博弈的扩展式 $\Gamma = \{N,H,P,u\}$ 如下：

（1）参与者集合：挑战者和在位者，$N = \{1,2\}$。

（2）全历史集合：$H = \{(Q),(E,A),(E,F)\}$。

（3）参与者函数：$P(\varnothing) = 1, P(E) = 2$。

（4）偏好：如果挑战者退出，那么挑战者得到 0 的利润，而在位者得到 3 的利润；如果挑战者进入，那么如在位者选择战斗，有可能两败俱伤，挑战者的利润为 -1，在位者的利润为 0，如果在位者选择和平共处，那么挑战者得 2 的利润，而在位者得 1 的利润。因而，挑战者偏好历史 $(E,A) > (Q) > (E,F)$，而在位者偏好历史 $(Q) > (E,A) > (E,F)$。其扩展式也可用博弈树来加以描述，如图 4-4 所示。

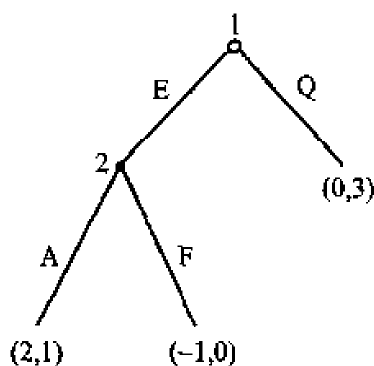


图 4-4　进入博弈

约定俗成，第一个数值表示参与者 1 的收益，第二个数值表示参与者 2 的收益，以此类推。需要注意的是，并不是所有的扩展式都能够用博弈树来表示，例如，当有 n 个参与者同时行动时，将难以用博弈树来表示。当参与者的行动不是离散的，而是连续的，或者博弈是无穷的，都无法用博弈树来表示。不过，如果一个动态博弈可以用博弈树表示，那么博弈树完全等价于该博弈的扩展式。由于博弈树直观、简便，因而几乎成为扩展式的唯一表现形式，最初扩展式的提出就是根据博弈树来定义的。在后面的分析中，我们也将用博弈树来表达动态博弈。

【例 4.4】 蜈蚣博弈

该博弈有两位参与者。当参与者 1 行动时，他将决定是结束博弈还是继续，如果结束博弈，那么参与者 1 得 2，参与者 2 得 0；如果继续博弈，那么轮到参与

者2决定是结束博弈还是继续,如果结束博弈,那么参与者1得3,参与者2得1;如果继续博弈,那么轮到参与者1行动,如果他选择左(L),那么参与者1得1,参与者2得2,如果他选择右(R),那么两人都得0。

该博弈的扩展式其博弈的扩展式 $\Gamma=\{N,H,P,u\}$ 如下:

(1) 参与者集合:$N=\{1,2\}$。

(2) 全历史集合:继续为C,结束为D,$H=\{(D),(C,D),(C,C,L),(C,C,R)\}$。

(3) 参与者函数:$P(\varnothing)=1$,$P(C)=2$,$P(C,C)=1$。

(4) 偏好:如果全历史为(D),那么参与者1得2,而参与者2得0;如果全历史为(C,D),那么参与者1得3,参与者2得1;如果全历史为(C,C,L),那么参与者1得1,参与者2得2;如果全历史为(C,C,R),那么两人都得0。参与者1最偏好历史(C,D),而参与者2最偏好历史(C,C,L)。该扩展式如图4-5所示。

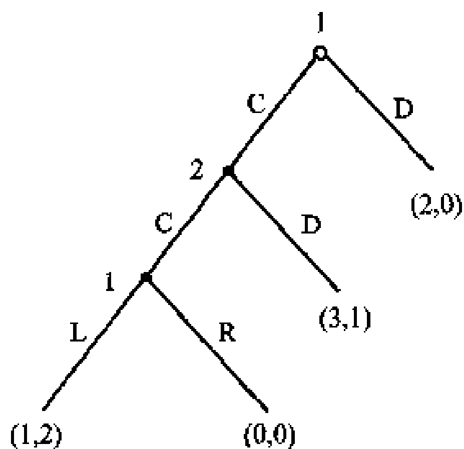


图4-5 蜈蚣博弈

㊓ 习题4.1 下象棋能否表示成一个扩展式,如果能,那么如何表达?下象棋(围棋也一样)虽然是最简单的动态博弈,但分析起来却不简单。

㊓ 习题4.2 请将生活中你所观察到的某些经济现象归纳成一个完全信息动态博弈模型,并用扩展式表达出来。

第三节 策略和结果

在策略式博弈中,策略是其三个组成要素之一,但在扩展式博弈中却没有策略的踪影,其原因就在于策略对于参与者而言,是一种"万全之策",而不再是单纯的行动。在静态博弈下策略等于参与者的行动或是行动的一个概率分布,它实际上并没有真正体现"万全之策"这个重要特征,只有在动态博弈(扩展式博弈)中策略的这一定义才能得到真正的体现。为了更加清楚地说明,我们用静态的囚徒困境与动态囚徒困境来做比较。

在静态囚徒困境中,我们已经知道囚徒1和囚徒2的策略等于行动,因而两个囚徒各有两个策略——招供和沉默。但在动态囚徒困境中,囚徒1和囚徒2的策略是不是仍然等于各自的行动呢?现在再次考察图4-1。囚徒1首先行动,他有两种行动可选,即招供或沉默,因而囚徒1的策略仍然为招供或沉默。

在囚徒 1 选择之后,囚徒 2 行动,他同样有两种行动可选,即招供或沉默。那么是不是囚徒 2 的策略也等于行动呢?如果囚徒 2 的策略仍然等于行动,就有出现失策的可能。因为囚徒 2 在囚徒 1 之后行动,策略等于行动将无法指明如果囚徒 1 一旦选择了某种行动(两种行动意味着有两种可能)囚徒 2 应当如何作出反应,这意味着囚徒 2 将不知所措,所谓的策略也就是假的,真正的策略应当能够针对囚徒 1 可能的各种选择作出相应的应对。由于囚徒 1 有两种行动,因而就有两种可能,每种可能下囚徒 2 有两种行动可以选择,因而,囚徒 2 的策略应该有四个,见表 4 - 1。

表 4 - 1　囚徒 2 的四个策略

	假如囚徒 1 选择招供	假如囚徒 1 选择沉默
策略 1	选择招供	选择招供
策略 2	选择招供	选择沉默
策略 3	选择沉默	选择招供
策略 4	选择沉默	选择沉默

每个策略都指明如果囚徒 1 采取了某种行动,那么相应地囚徒 2 应当如何去做。例如,策略 1 就指明如果囚徒 1 选择招供,那么就选择招供;如果囚徒 1 选择沉默,那么仍然选择招供。对于囚徒 2 而言,所谓策略就是一个完整的行动计划。

完整的行动计划比起单纯的行动,显然内涵丰富了许多,但对于真正的“万全之策”仍存在差距。所谓的“万全之策”不仅要考虑各种可能的情况,而且还应考虑到发生“意外”的可能,否则就不能称为万全之策。为了说明这一点,再来考察蜈蚣博弈。根据图 4 - 5,很容易发现参与者 2 的策略为 C 和 D,那么参与者 1 的策略又是什么呢?根据策略是一个行动计划的认识,我们可能认为参与者 1 的策略有三个,即 D(表示参与者 1 选择 D,然后游戏结束),CL(表示参与者 1 选择 C,如果参与者 2 选择 C,那么参与者 1 选择 L)和 CR(表示参与者 1 选择 C,如果参与者 2 选择 C,那么参与者 1 选择 R)。对于 CL 和 CR,相对容易理解,当参与者 1 选择 C 的时候,如果参与者 2 选择 D 那么游戏结束,因而 L 和 R 根本用不到,但参与者 1 在选择 C 的时候,根本不知道参与者 2 会如何做,因而他不得不给出参与者 2 选择 C 时自己如何行动。因而,CL 和 CR 就是参与者 1 的两个策略。而对于 D,就需要仔细斟酌了。当参与者 1 选择 D 时,游戏结束,根本轮不到参与者 2 选择,这是不是意味着参与者 2 根本不用作任何决策呢?显然这对参与者 2 而言太冒险了,“万一”参与者 1 没有选择 D,那怎么办?

因而尽管可能轮不到参与者2行动，但为了万无一失，参与者2仍应该给出一个策略以防万一。对于参与者1而言，同样应该考虑到这种"意外"，一旦参与者2作了决策，参与者1也应作出相应的决策。例如，不妨假设参与者1原本的策略为D，但实际（可能因为出错）选择了C，而参与者2也选择了C，那么这时参与者1将不知所措，因为参与者1的策略为D，它并没有指明如果发生了历史（C，C）这样的情况该如何做？这说明作为策略，行动计划D是不够的，DL和DR才是真正的策略。因而参与者1应该有四个策略：CL、CR、DL和DR，包含了所有可能和意外情况。对于DL或DR，也可以这样来理解："我打算选择D，但万一我出错选择了C，那么我将随后选择L或R。"所以说，策略是一个"万全之策"。

定义4.3 对于博弈$\Gamma=\{N,H,P,u\}$，参与者$P(h)$的一个**策略**$s_{P(h)}(h)$就是一个函数，它将每一个可能的历史h映射成行动空间$A_{P(h)}(h)$中的一个行动$a_{p(h)}$。

上述策略的定义实际上就是指当历史进行到某个阶段时，当轮到参与者i行动时，规定了他如何行动。例如，对于参与者1而言，一个策略就是当历史为空历史时，规定了参与者1如何行动，当历史为（C，C）时，规定了参与者1又如何行动，因而DL和DR就是参与者1的策略，至于历史（C，C）会不会发生那是另外一个问题，策略所要求的就是一旦出现了某个历史我应该如何做，而不能出现不知所措的情况。实际上，博弈越复杂，策略就越复杂，但根据定义4.3，参与者i只关心在每一个历史下（当轮到他行动时）如何做，其他都不重要，因而策略实际上又变简单了。

通过上面的说明可以看到，有什么样的策略组合就会有什么样的历史，但历史并不等于策略。为此，我们引入**结果函数**，即对于任意$s\in S$，存在某个$h\in H$，使得$O(s)=h$。参与者的收益函数u就是定义在结果上的函数。

例如，在蜈蚣博弈中，可知参与者1有四个策略CL、CR、DL和DR，参与者2有两个策略C和D，因而策略组合有8个，其相应的结果函数为

$$O(\mathrm{CL},\mathrm{C})=(\mathrm{CCL})\rightarrow u_1(O(\mathrm{CL},\mathrm{C}))=1 \text{ 和 } u_2(O(\mathrm{CL},\mathrm{C}))=2;$$

$$O(\mathrm{CR},\mathrm{C})=(\mathrm{CCR})\rightarrow u_1(O(\mathrm{CR},\mathrm{C}))=0 \text{ 和 } u_2(O(\mathrm{CR},\mathrm{C}))=0;$$

$$O(\mathrm{C}x,\mathrm{D})=(\mathrm{CD})\rightarrow u_1(O(\mathrm{C}x,\mathrm{D}))=3 \text{ 和 } u_2(O(\mathrm{C}x,\mathrm{D}))=1;$$

$$O(\mathrm{D}x,x)=(\mathrm{D})\rightarrow u_1(O(\mathrm{D}x,x))=2 \text{ 和 } u_2(O(\mathrm{D}x,x))=0。$$

其中x代表任意行动。上面的结果函数给了我们两点启示：一是要得到全历史实际上只需行动计划就可以了，不一定需要去考察所谓的"万全之策"，例如，$O(\mathrm{D},x)=\mathrm{D}=O(\mathrm{D}x,x)$是一样的，这样做的好处是能够简化分析，但在观念上，我们必须牢记策略是"万全之策"。二是图4-5的蜈蚣博弈实际上与

图4-6中的博弈完全等价，这就更为直观地指出了策略DL和DR的性质。实际上，汤普森（Thompson，1952）论证了对于任意两个等价的扩展式博弈，至少存在4种转换方式，通过转换，可以把复杂的扩展式博弈变成最简单的形式去分析。

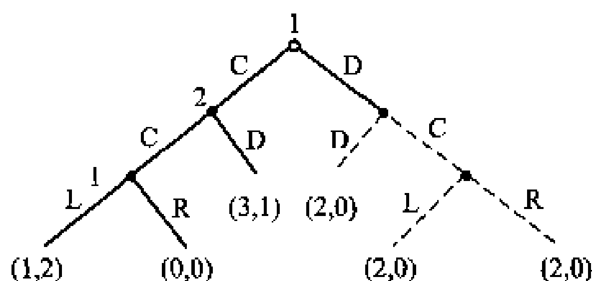


图4-6 与蜈蚣博弈等价的博弈

第四节 纳什均衡与子博弈完美均衡

由于动态博弈可以看作是由一串静态博弈组成，一个自然的想法是纳什均衡的概念能否在扩展式博弈中加以运用呢？答案是肯定的，但也带来了一些不理想的结果。由于纳什均衡的概念在动态博弈下要求过于宽松，因而一些不合理的均衡会在动态博弈中出现，而这些不合理的均衡大部分都是不可置信的（incredible）。下面将从纳什均衡入手，引出问题，最后给出子博弈完美均衡（subgame perfect equilibrium）的概念。

一、纳什均衡

纳什均衡概念的核心就在于，每一个参与者的策略都是给定其他参与者策略下的最优反应，并且对任意参与者成立。即便博弈是动态的，这一点也不会改变。那么，将静态博弈中的纳什均衡概念运用到动态博弈中应该是一个不错的思路，尽管这样做可能存在问题。

定义4.4 扩展式博弈$\Gamma=\{N,H,P,u\}$的纳什均衡是一个策略组合$s^*\in S$，使得对任一参与者$i\in N$，$\forall s_i\in S_i$，不等式$u_i(O(s_i^*,s_{-i}^*))\geqslant u_i(O(s_i,s_{-i}^*))$成立。

由于有什么样的策略就会有什么样的结果，因而收益函数又可直接看作是策略的函数，这样动态博弈$\Gamma=\{N,H,P,u\}$就有了与之等价的策略式博弈。

定义4.5 $G=\{N,S,u\}$称为完全信息动态博弈$\Gamma=\{N,H,P,u\}$的策略式（或基本式），其中S就是动态博弈$\Gamma=\{N,H,P,u\}$定义的策略空间，并且$u(s)=u(O(s))$。

为了加深理解，不妨考虑图4-7所示的完美动态博弈。

图 4－7 所示的完美动态博弈显然非常简单,读者应该很容易写出其扩展式。我们所要做的就是给出该动态博弈的策略式,其策略式 $G=\{N,S,u\}$ 如下:

(1) 参与者集合:$N=\{1,2\}$;

(2) 策略集合:$S=\{\{L,R\},\{c_1c_2,c_1d_2,d_1c_2,d_1d_2\}\}$;

(3) 偏好:$u_1(L,c_1c_2)=2,u_2(L,c_1c_2)=1$,其他依次类推。相对应的博弈矩阵如图 4－8 所示。

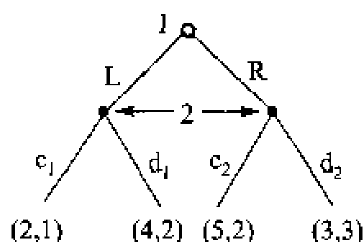


图 4－7 完美动态博弈

参与者 2

		c_1c_2	c_1d_2	d_1c_2	d_1d_2
参与者 1	L	2, 1	2, 1	4, 2	4, 2
	R	5, 2	3, 3	5, 2	3, 3

图 4－8 动态博弈的策略式

根据纳什均衡的定义,易知该动态博弈存在两个纳什均衡:(R,c_1d_2) 和 (L,d_1d_2)。对于均衡 (R,c_1d_2),只要我们稍加考察,就会发现这个纳什均衡含有不合理的因素,在现实中根本不会出现,原因就在于参与者 2 在历史(L)"威胁"出 c_1 是不可置信的,因为出 d_1 要比出 c_1 优($2>1$)。之所以出现这种情况,是由于当参与者 1 的策略为 R 时,历史进行到 L 的可能性为零,因此参与者 2 在历史 L 下无论采取什么行动都不会对他的最终收益造成影响。这意味着,纳什均衡这个概念对参与者 2 在不可能发生的历史 L 下如何选择并未作出规定,参与者 2 就有可能乱选(像一个非理性的人一样),而纳什均衡本身假设参与者是理性的,这就造成参与者 2 的策略是**动态不一致**的(dynamic inconsistent)。一个动态不一致的策略肯定不会是一个最优的策略。

我们也可以这样来理解参与者 2 的行动,参与者 2 之所以威胁当参与者 1 出 L 时,他要选择 c_1,目的在于通过威胁使参与者 1 选择有利于参与者 2 的 R,因为在参与者 1 选择 R 下,参与者 2 通过选择 d_2,能得到 3 的报酬,明显好于当参与者 1 选 L,参与者 2 选 d_1 时的收益 2。但我们要问的是,如果参与者 1 不顾参与者 2 的威胁而选择了 L,参与者 2 可能会出 c_1 吗?在参与者 2 为理性是公共信息的条件下,参与者 2 选择 c_1 的报酬为 1,而选择 d_1 的报酬为 2。由于 d_1 要优于 c_1,因而参与者 1 没有理由相信参与者 2 会实施他的威胁,也就是说,参与者 2 的策略 c_1d_2 是一个**不可置信的威胁**。如果威胁成真,c_1d_2 就是一个动态不一致的策略,因为参与者 2 事前是理性的,但在博弈进行到(L)时,他却成了一个非理性的人(选择了 c_1,而不是 d_1)。之所以出现上述问题,原因在于我们把

静态博弈中的纳什均衡概念生搬硬套到了动态博弈。一个纳什均衡只要求在博弈的总体上,参与者的策略须为均衡,而对博弈进行到某个部分时是否仍为均衡没有要求,这就可能导致总体和局部的冲突,产生一些不合理的结果。用数学的话来讲,这种情形之所以出现,是因为有些历史在原先的策略下发生的概率为0,因此在这些历史下所做的事情,对每个参与者的收益都没有影响,因而参与者在此可以任意选择行动而丝毫不会影响到原先纳什均衡的收益。对于策略组合$(L, d_1 d_2)$,就不存在这个问题,因为无论历史 R 是否发生,d_2对于参与者 2 而言都是最优的。因而,合理的纳什均衡应当是参与者 1 不理会参与者 2 的威胁,选择 L,参与者 2 在历史 L 选择 d_1,在历史 R 选择 d_2,因此真正的纳什均衡应该为$(L, d_1 d_2)$。

要消除动态博弈中的不可置信威胁,就需要一个比纳什均衡更强的均衡概念。它不仅在整个博弈中是均衡的,而且在局部也是均衡的;不但在现在是均衡的,在将来也应是均衡的。只有满足这个要求,博弈的参与者才能实现策略的动态一致性,这就导致了子博弈完美均衡概念的产生。

㊓ 习题 4.3　进入博弈、蜈蚣博弈的纳什均衡是什么?是否存在不可置信的威胁?

㊓ 习题 4.4　破釜沉舟。在两个国家之间存在一个小岛,双方都宣称对其拥有主权,小岛被其中一国的军队占据。另一国需要决定是否武力夺取小岛,对于守国来说,如果敌国来攻,要么通过船只撤退回国,要么誓死抵抗。对双方来讲,占据小岛而又不用打战最好,爆发战争结果最差。请你将其概括为一个完美信息动态博弈,并证明,如果守军炸沉船只,断了自己的退路,将会发生什么样的情况。

二、子博弈完美均衡

子博弈通俗地说就是动态博弈的局部,只要是一个真子历史,它后面的所有历史就构成了一个子博弈。如果一个动态博弈中的纳什均衡在所有该博弈的子博弈中也仍然是纳什均衡,那么前面遇到的不可置信威胁或是策略的动态不一致性就不会出现,这一思想最早由经济学家塞尔腾(Selten)在 1965 年提出[①]。也主要由于这个贡献,塞尔腾与纳什、豪尔绍尼共同获得了 1994 年的诺贝尔经济学奖。

定义 4.6　完全信息动态博弈 $\Gamma = \{N, H, P, u\}$ 的一个子博弈就是一个完整的动态博弈 $\Gamma(h) = \{N_s, H_s, P_s, u_s\}$,其具有如下特征:

① 1975 年塞尔腾用英语发表。

(1) 参与者集合 $N_s \subset N$；

(2) 给定真子历史 h,对于任意行动序列 h^*，$(h,h^*) \in H$，而$H_s = \{h^*\}$；

(3) $P_s = P$ 和 $u_s = u$。

在图 4-7 中,如果把整个博弈也看作是一个子博弈,那么有 3 个子博弈。历史 L 之后为第一个子博弈,历史 R 之后为第二个子博弈,加上原博弈,总共 3 个子博弈。在图 4-9中,存在五个子博弈,即原博弈 Γ、Γ(C)、Γ(D)、Γ(DE) 和 Γ(DF)。

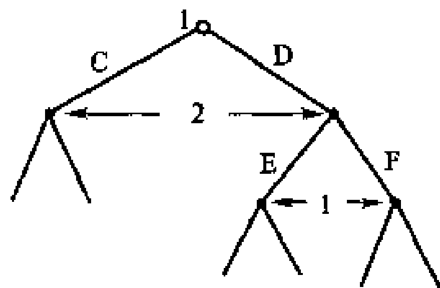


图 4-9　存在着五个子博弈

图 4-10 则给出了不是子博弈的情况。在图 4-10 中,虚线围起来的部分不是子博弈,因为它不构成一个完整的扩展式博弈。

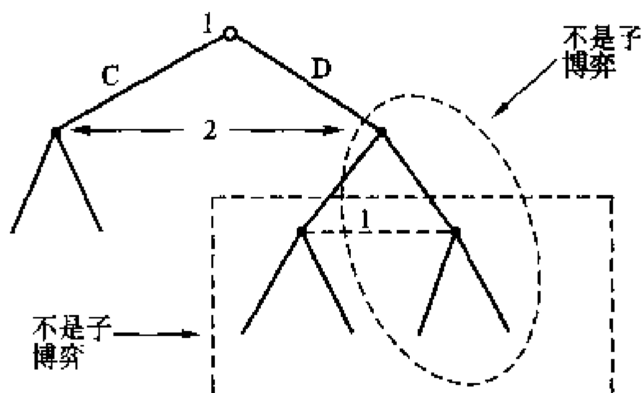


图 4-10　不是子博弈的例子

子博弈的核心就在于它本身就是一个博弈,因而纳什均衡的概念就可以运用到子博弈之中,这就使我们在前面建立起来的理论可以运用到动态博弈中,而没有必要另起炉灶。如果一个策略组合在所有的子博弈中都是纳什均衡,显然不可置信威胁和策略上的前后不一致都能被克服。

定义 4.7　在扩展式博弈 $\Gamma = \{N,H,P,u\}$ 中,如果一个策略组合 s^* 在所有该博弈的于博弈 $\Gamma(h) = \{N_s,H_s,P_s,u_s\}$ 中都是纳什均衡,那么我们就称策略组合 s^* 为子博弈完美均衡,即对任意参与者 $i \in P_s$ 和任意 s_i

$$u_i(O_h(s^*)) \geqslant u_i(O_h(s_i,s_{-i}^*))$$

成立。其中 h 为任意真子历史,$O_h(\cdot)$为子博弈 $\Gamma(h)$ 的结果函数。

根据定义 4.7,**可知子博弈完美均衡一定是纳什均衡,但反之不成立**。那么对于一个完全信息动态博弈我们如何去求解它的子博弈完美均衡呢? 一个一般的方法就是所谓的逆推法(backward induction)。

定义 4.8　所谓逆推法是指如下程序:

第一步，从扩展式博弈的终点开始，以找到该博弈的每一个最后子博弈（它不再包含任何其他更小的子博弈），然后求出纳什均衡，并计算出相应的收益。

第二步，将每一个最后子博弈的起点变成结束点，将计算出的每一个最后子博弈在纳什均衡下的收益写在其下方，我们就获得了一个新的扩展式博弈（或新的博弈树），称为压缩的扩展式博弈。这样经过一次压缩，就剔除了最后子博弈。

第三步，重复第一步和第二步，并重新得到一个压缩式博弈和相应的纳什均衡。这个过程一直进行到最后只剩下唯一一个子博弈为止，这时在逆推过程中找到的一系列子博弈的纳什均衡组合就是该扩展式博弈的一个完美均衡。

第四步，如果在逆推过程中没有遇到多重均衡，那么这个策略组合就是唯一的完美均衡；如果遇到了多重均衡，就需要对子博弈中每一个可能的均衡重复以上步骤，从而得出所有的完美均衡。比如，一个扩展式博弈有两个子博弈分别存在 2 个和 3 个纳什均衡，其他子博弈只有 1 个纳什均衡，那么该博弈就有 2×3=6 个完美均衡。

现在考虑一个假想的动态博弈，我们把它称为小蜈蚣博弈。假设有两个参与者 1 和 2，他们的扩展式博弈如图 4-11 所示。每一个正在作决策的参与者，都可以决定是继续进行游戏（C）还是终止游戏（D）。参与者 1 首先决定要停止游戏，还是继续下去。假如参与者 1 一开始就决定停止，则他的收益为 1，而对手的收益为 0。如参与者 1 决定继续游戏，则轮到参与者 2 作出决定。如果他选择停止，他的收益为 3，参与者 1 的收益为 -1。如果他选择继续，则又轮到参与者 1 作出决定，如此等等。显然小蜈蚣博弈包含许多子博弈，如果从正面来求解小蜈蚣博弈的完美均衡会非常麻烦，特别是小蜈蚣博弈扩展到 200 次、300 次更是如此，但是用逆推法则非常简单。

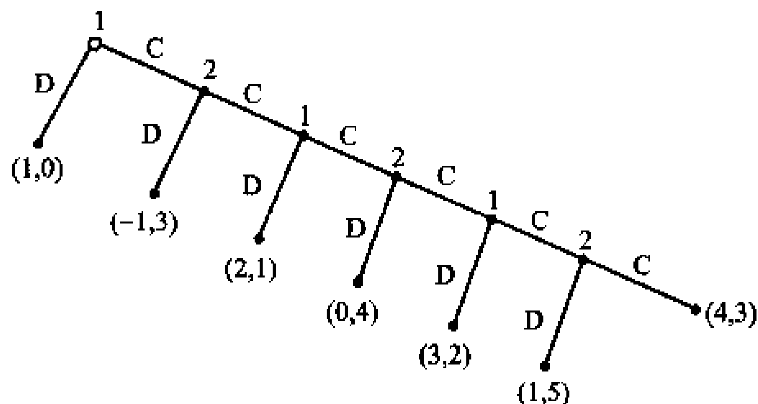


图 4-11 小蜈蚣博弈（1）

首先求最后子博弈的纳什均衡。在最后子博弈,参与者 2 决定是选择 C 还是选择 D。如果选择 C,参与者 2 的收益为 3;如果选择 D,参与者 2 的收益为 5,所以参与者 2 在倒数第一个子博弈的纳什均衡为 D。参与者 1 的收益为 1。现在我们将两人的收益代人倒数第二个子博弈。其如图 4 - 12 所示。

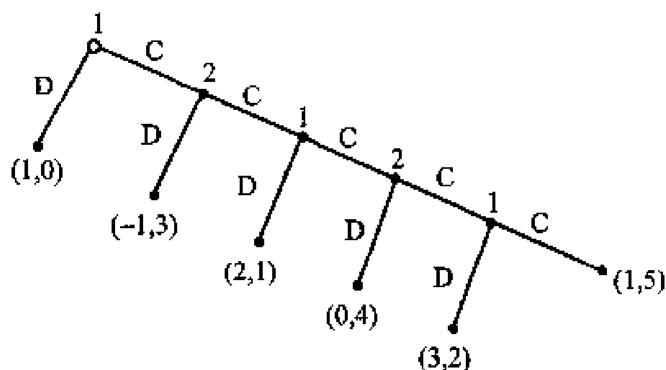


图 4 - 12 小蜈蚣博弈(2)

现在再考虑图 4 - 12 中的最后子博弈。如果参与者 1 选择 C,那么他的收益为 1;如果选择 D,收益为 3。所以,纳什均衡为参与者 1 选择 D,收益向量为(3,2)。再将它代人倒数第三个子博弈。如此不断重复,一直到最后只剩下唯一一个子博弈为止。如图4 - 13。这时参与者 1 的最优选择为 D。

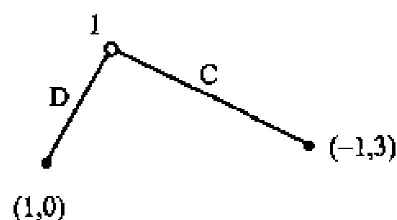


图 4 - 13 小蜈蚣博弈(3)

这样不断地往回解,我们可以知道任何参与者在面临选择时,都会选择结束游戏 D。因而小蜈蚣博弈唯一的完美均衡为:参与者 1 一开始就选择 D 结束游戏,如果不是参与者 2 将选择结束 D,如果不是参与者 1 将选择结束 D,如此等等,即每一个参与者都选择 D。

实际上逆推法只不过是数学中的动态规划在扩展式博弈中的运用而已,根据动态规划理论,只要动态的阶段是有限的,最优规划(均衡)必然存在。

定理 4.1 只要扩展式博弈 $\Gamma = \{N,H,P,u\}$ 是有限的,即参与者有限、行动空间有限、博弈的阶段有限(不是无穷进行下去),那么扩展式博弈 Γ 至少存在一个子博弈完美均衡。

因为子博弈完美均衡是从最后一期的子博弈开始,而最后子博弈就好比是一个静态博弈,由于参与者和可用行动是有限的,因而子博弈必然存在纳什均衡(纳什定理),然后压缩博弈,由于博弈的阶段有限,所以最后总能把完全信息扩展式博弈压缩成一个静态博弈,这就意味着扩展式博弈存在着子博弈完美均衡。在前面我们证明纳什定理的过程中,所谓行动空间有限,实际上就是要求行动空

间是紧致的，这里也一样，因而将行动空间有限改称行动空间紧致可能更为准确。

那么在运用逆推法的过程中，会不会遗漏掉某个子博弈完美均衡呢？

定理 4.2 等价性定理 $\{s^*\}$为有限扩展式博弈 $\Gamma=\{N,H,P,u\}$ 的所有子博弈完美均衡的集合，$\{s^\#\}$为该扩展式博弈运用逆推法找到的所有子博弈完美均衡的集合，那么$\{s^*\}=\{s^\#\}$，即子博弈完美均衡与逆推法是完全等价的。

【例 4.5】 一个简单博弈的子博弈完美均衡

图 4－14 是一个相当简单的扩展式博弈，但却存在着 6 个子博弈完美均衡，要直接找出这 6 个子博弈完美均衡并不容易，而运用逆推法就简单得多。

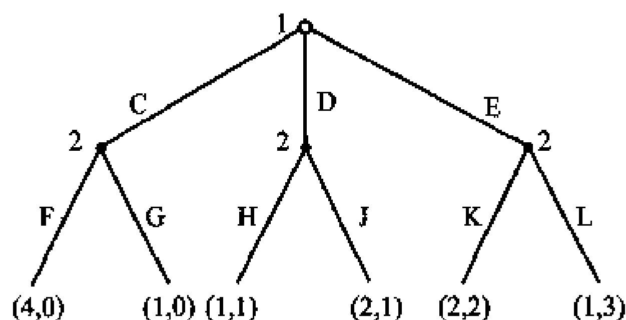


图 4－14 一个简单博弈

运用逆推法，可得子博弈 Γ(C)的纳什均衡为 F 和 G，子博弈 Γ(D)的纳什均衡为 H 和 J，子博弈 Γ(E)的纳什均衡为 L，由此可得参与者 2 的 4 个均衡策略：(FHL)，(FJL)，(GHL)和(GJL)。现在，回到原博弈，寻找原博弈的纳什均衡。分四种情况：① 给定参与者 2 的策略(FHL)，参与者 1 的最优策略为 C，因而(C，FHL)为完美均衡。② 给定参与者 2 的策略(FJL)，参与者 1 的最优策略为 C，因而(C，FJL)为完美均衡。③ 给定参与者 2 的策略(GHL)，参与者 1 的最优策略有三个：C、D 和 E。因而，完美均衡也有三个：(C，GHL)，(D，GHL)和(E，GHL)。④ 给定参与者 2 的策略(GJL)，参与者 1 的最优策略为：D。因而，完美均衡为：(D，GJL)。综上所述，即该博弈存在 6 个完美均衡。

从上面这个例子，可以得出这样一个判断，即如果每一个全历史的收益都不相等，那么扩展式博弈一定存在唯一的子博弈完美均衡。

命题 4.1 在扩展式博弈 $\Gamma=\{N,H,P,u\}$ 中，如果每一个全历史对应的参与者的收益都不相等，那么存在唯一的子博弈完美均衡。

逆推法虽然在求解完美均衡上非常有效，但也有缺陷，当博弈为无穷动态博弈时，我们将无法运用逆推法。

(易) 习题 4.5 团队合作。两人需要共同合作来完成一项任务，第一个人首先选择努力水平 a_1，然后第二个人选择努力水平 a_2，努力水平为非负数。他们的收益函数为 $a_i(c+a_j-a_i)$，$i,j=1,2,i\neq j,c>0$。该博弈的扩展式和子博弈完

美均衡是什么？与例 2.15 协同作业相比有什么差别？

㊟ 习题 4.6　人民币拍卖。有价值为 v 的人民币进行拍卖，v 为元的整数，两人进行竞拍。拍卖的规则是这样的，两人轮流竞拍，竞拍价必须为元的整数，轮到某人竞拍时，他可以"过"，也可以给出一个更高价。如果他"过"，那么对手得到 v 的人民币，两人都必须支付最后报出的竞拍价。例如，第一个人开始竞拍，如果他"过"，那么竞拍结束，第二个人得到人民币，两人都支付 0 元。如果第一个人报 $b_1>0$，第二个人报 $b_2>b_1$，接下来如果第一个人"过"，那么第二个人得到人民币，支付 b_2，第一个人支付 b_1。两人都拥有 $w(>v)$ 的财产，两人的报价都不得超过其拥有的财产。将其概括为扩展式博弈，并找到其子博弈完美均衡。

第五节　经典举例

在这一节中，我们将运用前面所学的理论来具体分析一些经典模型。

一、价格领先制

价格领先制是垄断厂商非公开勾结的一种普遍形式，最早由施塔克尔贝格(Stackelberg)提出，因而又被称为施塔克尔贝格模型。这里假设市场存在两个垄断企业——领先厂商(leader)和跟随厂商(follower)，领先厂商决定一个价格 p，随后跟随厂商选择产量。市场的需求函数为 $Q(p)$，其中 $Q=q_L+q_F$ 为市场总需求，q_L 为领先厂商的产量，q_F 则为跟随厂商的产量。设领先厂商的成本函数为 $C_L(q)$，跟随厂商的成本函数为 $C_F(q)$，$C_F(q)>C_L(q)$，两个成本函数都满足 $C'(q)>0$，$C''(q)>0$。该博弈的扩展式 $\Gamma=\{N,H,P,u\}$ 如下：

(1) 参与者集合：$N=\{L,F\}$，L 代表领先厂商，F 代表跟随厂商；

(2) 全历史集合：$H=\{p,q_L\}$，$p\geqslant 0$；$q_L\geqslant 0$；

(3) 参与者函数：$P(\varnothing)=1$，$P(p)=2$；

(4) 偏好：垄断企业的偏好可用各自的利润函数表示。

我们运用逆推法来解这个动态博弈的完美均衡。首先计算跟随厂商在给定 p^* 下的最优产量 $q_F(p)$，即

$$\max_{q_F} p^* q_F - C_F(q_F)$$

一阶条件为

$$p^* = C_F'(q_F) \tag{4-J-1}$$

即价格必须等于边际成本。由于边际成本函数 C_F' 为一一对应的连续函数，所以

存在反函数，即 $q_F(p^*) = MC_F^{-1}(p^*)$，$MC_F$为跟随厂商的边际成本。

由于领先厂商也能够像跟随厂商一样解出跟随厂商的最优反应，领先厂商就可以预测到他如选择 p^*，跟随厂商将根据 $MC_F^{-1}(p^*)$来选择产量 q_F。因而，在博弈的第一阶段，领先厂商的问题就可表述为

$$\max_p p[Q(p) - q_F(p)] - C_L(Q(p) - q_F(p))$$

解一阶条件得

$$p = C_L'(q_L) - \frac{q_L}{Q' - q_F'} \quad (4-J-2)$$

由于 $Q' < 0$（需求定理），$q_F' > 0$（供给定理），所以上式一定大于0。根据导数公式 $\frac{\mathrm{d}f^{-1}(x)}{\mathrm{d}x} = \frac{1}{f'(x)}$，上式又等于

$$p = C_L'(q_L) - \frac{q_L}{Q'(p) - \frac{1}{C_F''(q_F)}} \quad (4-J-3)$$

(4-J-3)式实际上是一个隐函数，要得出明确的解有一定难度，不过根据假设，仍能得出一些认识：当 q_F越接近边际成本的最低点（回忆边际成本曲线为U型，停止营业点为可变平均成本的最低点，它最接近边际成本最低点），$C_F''(q_F) \to 0$，从而 $p \to C'(q_L)$，即领导厂商的定价如果越接近跟随厂商的停止营业点，跟随厂商的产量就越小，从而领导厂商的市场份额就越大，当领导厂商的定价等于跟随厂商的停止营业点（因为 $C_F(q) > C_L(q)$，所以存在这种可能），那么跟随厂商的产量为零，而领导厂商占有整个市场，所以它的定价趋于边际成本，即 $p \to C_L'(q_L)$。因而[(4-J-2)，(4-J-1)]就构成了价格领先制的完美均衡。例如，假设 $C_F(q_F) = q^2/2$，$C_L(q_L) = cq_L$，$Q(p) = a - bp$，可知 $C_F'(q_F) = q_F$，$C_F''(q_F) = 1$，$Q'(p) = b$，$C_L'(q_L) = c$，因而(4-J-3)式等于

$$p = c + \frac{q_L}{b+1}$$

已知 $q_L = a - (b+1)p$，所以上式又等于

$$p = \frac{a + c(b+1)}{2(b+1)}$$

因而$\left(p = \frac{a + c(b+1)}{2(b+1)}, q_F = p\right)$就是该博弈的于博弈完美均衡。

如果我们把上述模型稍加改造——领先厂商不是选择价格而是选择产量，会得出一个有趣的现象，即领先厂商占有更多的信息反而带来了不利，而这种情况在静态博弈中不会出现。

(易) 习题4.7　产量领先制。领先厂商不是选择价格，而是选择产量。请你

写出其扩展式，并与价格领先制的完美均衡进行比较。

二、收买选票

一个拥有 k 个议员的议会对两个议案 B_1 和 B_2 进行投票表决，议会议员不能弃权，只能投一次票，要么支持 B_1，要么支持 B_2，获得过半票数的议案将获得通过。假设存在两个利益集团，如果 B_1 获得通过，那么将给利益集团 1 和利益集团 2 分别带来 v_1 和 0 的收益；如果 B_2 获得通过，那么将给利益集团 1 和利益集团 2 分别带来 0 和 v_2 的收益。利益集团可以向议会议员购买选票，议员则根据谁给的钱多投谁的票，如果两个利益集团给的钱一样，那么将把选票投给利益集团 2。为了让有利于自己的议案通过，利益集团 1 首先行动，即向议员收买选票，为 $x=(x_1,\cdots,x_k)$，随后利益集团 2 行动，为 $y=(y_1,\cdots,y_k)$，其中 x,y 都是非负实数向量，并且 $x_1\leqslant x_2\leqslant\cdots\leqslant x_k$。这里我们分析 $k=2n+1$ 的情况（n 为自然数）。该博弈的扩展式 $\Gamma=\{N,H,P,u\}$ 为：

（1）参与者集合：$N=\{1,2\}$；

（2）全历史集合：$H=\{x,y\}$；

（3）参与者函数：$P(\varnothing)=1,P(x)=2$；

（4）偏好：利益集团 1 的收益函数为 $u_1=v_1-x_1-\cdots-x_k$（当 B_1 通过）或 $u_1=-x_1-\cdots-x_k$（当 B_2 通过），利益集团 2 的收益函数同理。

那么该博弈的子博弈完美均衡是什么呢？

运用逆推法。首先，考虑利益集团 2 的纳什均衡策略。给定利益集团 1 的行动 x，那么利益集团 2 在选择最优策略时应当考虑三种情况：一种是有利可图，这时第 2 利益集团应收买前 $(k+1)/2=m$ 个议员投 B_2 的票（多了要花更多的钱，少了不过半），我们把这 m 个议员从利益集团 1 所得货币之和定义为 m_x，要有利可图就意味着 $v_2>m_x$，这时利益集团 2 的最优行动为 $y=(x_1,\cdots,x_m,0,\cdots,0)$，即对前 m 个议员给与利益集团 1 一样的钱，其他则不给。第二种是无利可图，这时 $v_2<m_x$，因而 $y=(0,\cdots,0)$，即不给任何钱。第三种是 $v_2=m_x$，这时利益集团 2 选 $y=(x_1,\cdots,x_m,0,\cdots,0)$ 和 $y=(0,\cdots,0)$ 收益都一样，因而无所谓。

给定利益集团 2 的最优策略，那么利益集团 1 又应当如何选择呢？显然，如果利益集团 1 对于前 m 个议员给出的钱 $m_x<v_2$，那么 B_1 肯定不能通过，所以利益集团 1 最优的策略是 $x=0$（表示向量的每一个元素都等于 0）。如果 $m_x>v_2$，那么 B_1 肯定通过，所以利益集团 1 的最优策略就是使得给每一个议员的钱 x_i 都等于 m_x/m（实际上，利益集团 1 给每一个议员一样的钱比不一样要严格优）。如果 $m_x=v_2$，那么 B_1 在利益集团 2 选择 $y=(x_1,\cdots,x_m,0,\cdots,0)$ 时不通过，在利益集团 2 选择 $y=0$ 时通过。对于利益集团 1 而言，是否有利可图关键就在

于 v_1 的值。由于利益集团 1 给每一个议员的钱一样比不一样要严格优，所以当 $v_1 > km_x/m$，并且 $m_x \geqslant v_2$，也就是 $v_1 > kv_2/m$ 时，那么利益集团 1 的最优决策为 $x = m_x/m$，由于货币是连续可分的，所以 $m_x = v_2$。如果 $v_1 < kv_2/m$，硬要让 B_1 通过，只会使得 $v_1 - km_x/m < 0$，其中 $m_x \geqslant v_2$，所以不如选择 $x = 0$。如果 $v_1 = kv_2/m$，那么情况就有点复杂。要让 B_1 通过除非 $m_x = v_2$，这时 $x = 0$ 与 $x = v_2/m$ 对于利益集团 1 都一样。如果 $x = 0$，那么 $y = 0$ 显然是子博弈完美均衡。如果 $x = v_2/m$，要么 $y = 0$，要么 $y = (v_2/m, \cdots, v_2/m, 0, \cdots, 0)$。如果是 $y = (v_2/m, \cdots, v_2/m, 0, \cdots, 0)$，那么将使 B_2 通过，所以对于利益集团 1 而言，它有动机改变 $x = v_2/m$，例如，给每一个议员的钱增加一个 $\Delta \to 0$，将使 B_1 通过，付出的代价为 $v_1 - x_1 - \cdots - x_k - k\Delta$，显然好于 $-x_1 - \cdots - x_k$。实际上，无论是哪种情况，$[x = v_2/m, y = (v_2/m, \cdots, v_2/m, 0, \cdots, 0)]$ 都不可能是子博弈完美均衡。

综上所述，收买选票的子博弈完美均衡为：

如 $v_1 > 2kv_2/(k+1)$，那么利益集团 1 选择 $x = 2v_2/(k+1)$，利益集团 2 选择 $y = 0$；

如 $v_1 < 2kv_2/(k+1)$，那么利益集团 1 选择 $x = 0$，利益集团 2 选择 $y = 0$；

如 $v_1 = 2kv_2/(k+1)$，那么利益集团 1 选择 $x = 0$，利益集团 2 选择 $y = 0$；或者 $x = 2v_2/(k+1)$，利益集团 2 选择 $y = 0$。

收买选票博弈最显著的特征就是：无论什么情况，子博弈完美均衡中利益集团 2 的最优策略都是 0，即只要存在利益集团 2 也会收买选票，就有可能使利益集团 2 不用付出任何代价就能获得最大收益。因而，像产量领先制一样，先行动者总是不利，而后行动者总是具有优势，我们把这种现象称为后发优势。

㊥ 习题 4.8　稍加观察就能发现各国议会中的议员数都是奇数，现在请你分析 $k = 2n$ 的情况。（提示：需要对投票出现势均力敌的情况作出假设）

㊦ 习题 4.9　将第三章中的多党竞选稍作修改：设党派为 3，各党派依次选择其政治主张或是退出竞选，对于每个党派而言赢是最好的，其次为打个平手，如果失败，那还不如退出竞选，请分析该博弈。

三、赛跑

前面分析的例子具有后发优势，那么存不存在先发优势呢？这里我们介绍赛跑博弈，赛跑博弈可以运用在各种各样的具有先发优势的竞争模型中。

选手 1 和选手 2 赛跑，他们离终点线的距离分别用 k_1 和 k_2 来表示，k 表示距离终点线的步数。两个选手轮流行动，选手的行动可以是不动，也可以移动 1 步或是 2 步，成本分别为 $c(0) < c(1) < c(2)$。如果选手 1 先到达终点，那么得到 v 的奖励，净收益为 $v_1 - c_1$，而选手 2 的收益就为 $-c_2$，c 为每阶段选手走过的步数

的成本之和，例如，选手 1 走过的步数为 0,1,2，那么 $c_1 = c(0) + c(1) + c(2)$。对于选手 2 同理。如果两个选手都没有移动，那么游戏结束。该博弈的扩展式 $\Gamma = \{N, H, P, u\}$ 如下：

(1) 参与者集合：$N = \{1,2\}$；

(2) 全历史集合：$H = \{(x^1, y^1, x^2, y^2, \cdots, x^m, y^m)\}$，其中 x 为选手 1 的步数，y 为选手 2 的步数，$x, y \in \{0,1,2\}$，$m \leqslant k_1, m \leqslant k_2$。

(3) 参与者函数：$P(\varnothing) = 1, P(\cdots, x) = 2, P(\cdots, y) = 1$；

(4) 偏好：每个选手都希望赢得奖品 v，同时收益为正，否则选择不移动。

我们可将不同的赛跑博弈表示成 $\Gamma(k_1, k_2)$。为便于说明，不妨假设 $7 > v > 6, k_1 = k_2 = 6, c(x) = x^2$，那么赛跑博弈 $\Gamma(6,6)$ 的子博弈完美均衡是什么？正面求解显得十分繁琐，不妨运用逆推法。

我们首先考虑 $\Gamma(1,1)$ 的情况，即两位选手离终点线只差一步。在这种状态下，谁先行动谁赢，移动一步将能收获 6 而付出 1，因而移动一步比不移动要优，而后行动者将不移动（移动的话同样不能赢，但却要多付出 1）。当情况为 $\Gamma(1,2)$ 时，谁先行动谁赢。如果是选手 1 先行动，他移动一步将得 6 而付出 1，移动一步比不移动要优，而选手 2 不移动；如果是选手 2 先行动，那么他移动两步将得 6 而付出 4，移动两步比移动一步或不移动要优，而选手 1 不移动。当情况为 $\Gamma(2,1)$ 时，分析与 $\Gamma(1,2)$ 完全一样。当情况为 $\Gamma(2,2)$ 时，谁先行动谁赢，如果选手 1 先行动，选手 1 移动两步比移动一步或不移动要优，而选手 2 不移动。因而，上述四种情况都意味着，谁先行动谁赢，而后行动者不会移动。

接下来考虑 $\Gamma(1, k \geqslant 3)$ 的情况，即使选手 2 首先行动，如果他走一步，那么博弈变为 $\Gamma(1, k \geqslant 2)$；如果他走两步，那么博弈变为 $\Gamma(1, k \geqslant 1)$。这都意味着选手 1 赢，因而选手 2 不会移动，而选手 1 必会移动一步。同理，$\Gamma(k \geqslant 3, 1)$ 也是一样。当情况为 $\Gamma(2, k \geqslant 3)$ 时，即使选手 2 首先行动，如果他走一步，那么博弈变为 $\Gamma(2, k \geqslant 2)$；如果他走两步，那么博弈变为 $\Gamma(2, k \geqslant 1)$。这都意味着选手 1 会走两步赢得比赛，因而选手 2 不会移动，而选手 1 必会移动一步。同理，$\Gamma(k \geqslant 3, 2)$ 也是一样。

其次考虑 $\Gamma(3,3)$ 的情况，显然如果选手 1 先行动（走一步比走两步要好），那么博弈将变为 $\Gamma(2,3)$；如果选手 2 先行动，那么博弈将变为 $\Gamma(3,2)$，无论哪种情况都意味着谁先行动谁赢，先行动者每次移动一步，后行动者不移动。当情况为 $\Gamma(3,4)$ 时，如果选手 1 先行动（走一步比走两步要好），那么博弈变为 $\Gamma(2,4)$，选手 1 肯定赢。所以选手 1 每次走一步，选手 2 不移动。如果选手 2 先行动，那么他会走两步（因为走一步将使博弈变为 $\Gamma(3,3)$，轮到选手 1 行动，所以选手 2 必输），博弈变为 $\Gamma(3,2)$，选手 2 会赢，所以选手 2 走两步，然后走两

次一步，选手 1 不移动。同理，$\Gamma(4,3)$分析完全一样。当情况为 $\Gamma(4,4)$时，如果选手 1 先走，他必走两步(因为走一步，博弈将变为 $\Gamma(3,4)$，这时轮到选手 2 走，所以选手 1 必输)，这时博弈变为 $\Gamma(2,4)$，选手 2 的最优策略是不移动，然后选手 1 走两次一步到达终点。总之，上述四种情况都意味着谁先走谁赢，后行动者不移动。$\Gamma(k\geq 5,3)$和 $\Gamma(k\geq 5,4)$同理，不再赘述。实际上，当距离终点线不低于 5 步时，选手每次只能走一步，否则收益将为负，因而要么每次走一步，要么一步不走。

最后考虑 $\Gamma(5,5)$和 $\Gamma(6,6)$的情况，那么谁先走一步选赢，后行动者不移动。当情况为 $\Gamma(5,6)$和 $\Gamma(6,5)$时，将是距离终点线最近的选手每次移动一步，另一位选手不移动。

赛跑博弈整个分析的结果如图 4－15 所示。图 4－15 中的 f 表示谁先行动谁赢，1 表示无论谁先移动，选手 1 都会赢，2 同理，实箭头表示选手 1 需要先走两步才能赢，虚箭头则表示选手 2 需要先走两步才能赢。

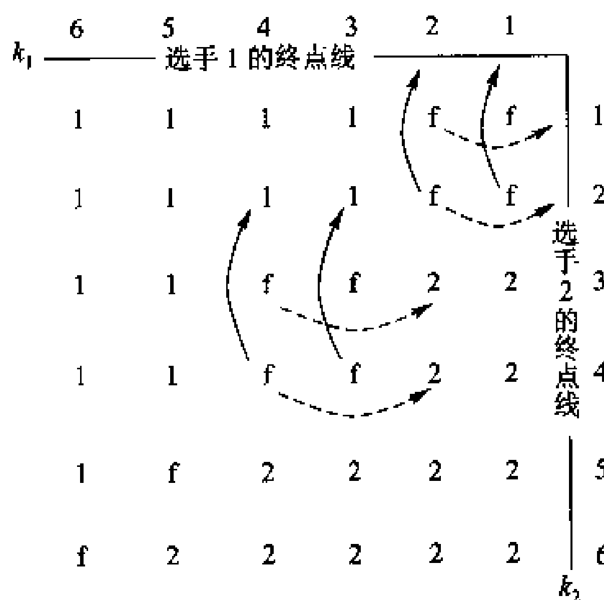


图 4－15 赛跑的子博弈完美结果

实际上，只有满足条件 $k_1\leq v/c(1)$和 $k_2\leq v/c(1)$，赛跑博弈才有意义，否则如果选手每次移动一步，达到终点线所费成本大于 v，那么还不如不移动。

(易) 习题 4.10 请你总结 $\Gamma(6,6)$的子博弈完美均衡。

(中) 习题 4.11 (赛跑游戏的一般化)请你考虑 $n+1>v>n$，n,k_1,k_2都为任意自然数的情况。

四、工会与企业博弈

在一个劳动力市场上存在一个垄断性质的工会，工会首先决定工资率，而

后企业根据工资来决定需要雇佣多少工人。工会的效用函数为 $U(w,L)$，其中 w 为工会向企业开出的工资水平，L 为就业量。假定 $U(w,L)$ 是 w 和 L 的增函数，即工资率越高，就业量越高，工会取得的成绩越大。企业的利润函数为 $\pi(w,L) = R(L) - wL$，其中 $R(L)$ 为企业雇佣 L 名工人在最优的生产和产品市场决策下可获得的收入，假定 $R(L)$ 为凹函数，即 $R'(L) > 0$，$R''(L) < 0$，如图 4－16 所示。

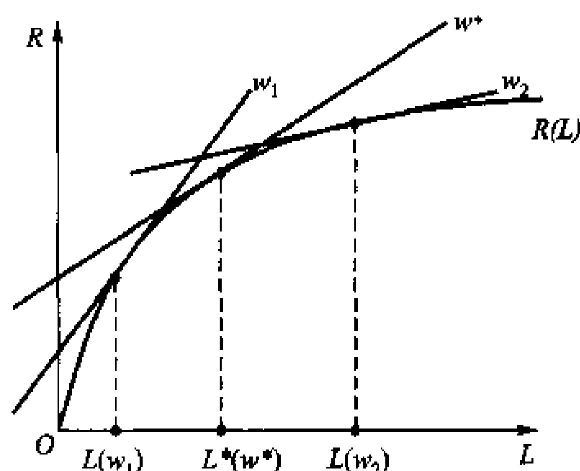


图 4－16 凹函数

该博弈的扩展式 $\Gamma = \{N,H,P,u\}$ 如下：

(1) 参与者集合：N = {工会，企业}；

(2) 全历史集合：$H = \{w,L\}$；

(3) 参与者函数：$P(\varnothing)$ = 工会，$P(w)$ = 企业；

(4) 偏好：工会和企业的收益函数分别为效用函数 $U(w,L)$ 以及利润函数 $\pi(w,L) = R(L) - wL$。

我们仍然利用逆推法来求解该动态博弈的子博弈完美均衡。

首先，在给定工会工资率 w，企业应选择 $L(w)$ 以使利润最大化

$$\max_{L} R(L) - wL$$

一阶条件为

$$R'(L^*) - w^* = 0$$

由于 $R(L)$ 为凹函数，上式必有解，并且一阶条件也是充分条件。利用一阶条件的反函数，就能得到 $L(w)$，它实际上表示在利润最大化下，企业如何根据工资率 w 来选择 L。根据图 4－16 可知，当 $L>L^*$ 时，$R'(L) - w^* < 0$（这表明当 L 保持不变时，工资率 w 降低会使企业的利润提高）；当 $L<L^*$ 时，$R'(L) - w^* > 0$，所以企业的等利润曲线是一个倒 U 型，$L(w)$ 必定过各等利润曲线的最高点，较低的等利润曲线代表较高的利润水平，如图 4－17 所示。

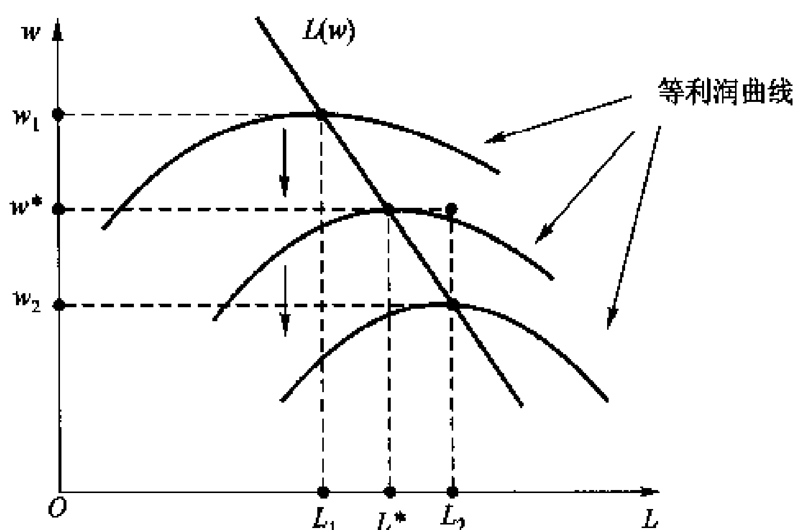


图 4－17　企业的等利润曲线

最后我们分析工会的最大化问题。工会在第一阶段的问题是根据企业的决策 $L(w)$，选择工资率以使工会的效用函数最大化，即

$$\max_{w} U(w, L(w))$$

由于我们不知道效用函数的具体形式，因而我们用工会的无差异曲线来代替上述公式最大化。根据微观经济学的知识，当无差异曲线与预算线相切时，切点所代表的组合将使效用最大化。在上面的分析中，图 4－16 中的 $L(w)$ 就是工会面临的预算线，因为企业追求利润最大化必然导致这一结果。在图 4－18 中，工会的最优选择就是使得它的无差异曲线与 $L(w)$ 曲线相切。从面，(w^*, L^*) 就是工会和企业博弈的完美均衡。

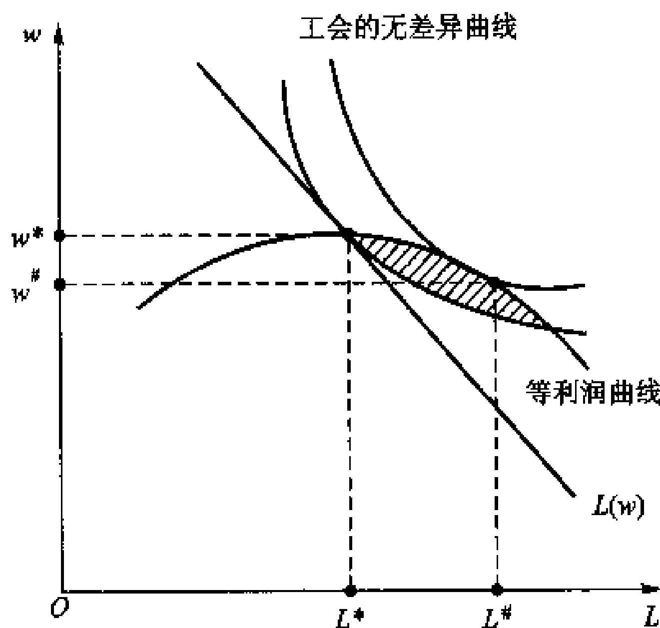


图 4－18　最优选择与无效率

但这个子博弈完美均衡不满足帕累托效率，因为如果 w 和 L 位于阴影部分之内，则工会和企业的福利都能得到改善。例如，(w^{s},L^{s}) 在保证企业利润不变的情况下，使得工会的收益高于完美均衡 (w^{*},L^{*}) 下的收益，因而存在帕累托改善。上述分析的结论成为批评工会控制工资率导致无效率的一个重要理由。在第五章重复博弈中，我们将看到如果博弈无限期地重复进行（或不知道博弈什么时候结束），那么完美均衡就有可能在阴影部分之内出现。

㊂ 习题 4.12　给定企业的生产函数为 $K^{a}L^{1-a}$，r 为资本 K 的价格，w 为劳动 L 的价格，产品市场完全竞争。工会的效用函数为 $w^{b}L^{1-b}$，$a,b\in(0,1)$，其他条件不变，那么工会和企业博弈的子博弈完美均衡是什么？

以上例子都是完美信息动态博弈，下面的例子则是非完美信息博弈。

五、开放式基金赎回潮

中国 2006 年最热的恐怕就属"基金热"了，但基金的存在却可能加大股市的波动。通常，当股市出现大幅下跌时，基金购买者为了保住基金净值，有可能出现大面积的基金赎回，这反过来又进一步加大股市的下跌。这里我们分析一个简单的基金赎回潮模型。假设有两人投资 D 元购买了一家基金公司的开放式基金，基金公司将这些钱投入到股市中，只有在长期内，基金才能有收益。为了简化，我们假设没有贴现。在短期，两个投资人可以赎回，也可以不赎回，如果赎回意味着基金公司被迫卖出股票，能收回的钱为 $2r$，这里 $D>r>D/2$。如果只有一个人在短期赎回，那么他得到 D，另一个人得到 $2r-D$。如果两个人都赎回，各自得到 r。如果两人都不赎回，表示长期投资。在长期，如果两人都赎回，那么每人得 R，其中 $R>D$（连本带利）。如果只有一人赎回，那么赎回人得 $2R-D$，不赎回人得 D。如果两人都不赎回，则基金公司向每个投资者返还 R。

这个博弈的扩展式 $\Gamma=\{N,H,P,u\}$ 如下：

(1) 参与者集合：$N=\{1,2\}$。

(2) 全历史集合：设 W 表示赎回，S 表示不赎回，$H=\{$(W,W)，(W,S)，(S,W)，[(S,S)，(W,W)]，[(S,S)，(W,S)]，[(S,S)，(S,W)]，[(S,S)，(S,S)]$\}$。

(3) 参与者函数：$P(\varnothing)=\{1,2\}$，$P(S,S)=\{1,2\}$。

(4) 偏好：对于任一投资人而言，收益越大越好，相应收益如图 4－19。

很明显，这个模型存在两个子博弈。根据逆推法，我们先从最后一个子博弈开始分析。显然，在长期内，对于任一投资人而言，无论对手如何行动，选择赎回都是最优策略，即赎回是投资人的严格优策略。因而最后一个子博弈的纳什均衡为（赎回，赎回），相应的收益为 (R,R)。利用上述结论，将其结果带入到短期，就得到图 4－20。

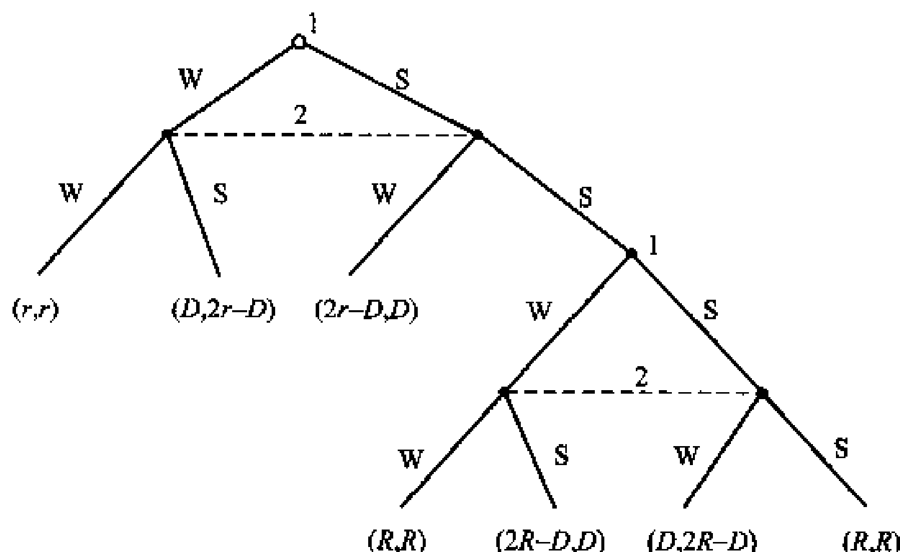


图 4 - 19　基金赎回(1)

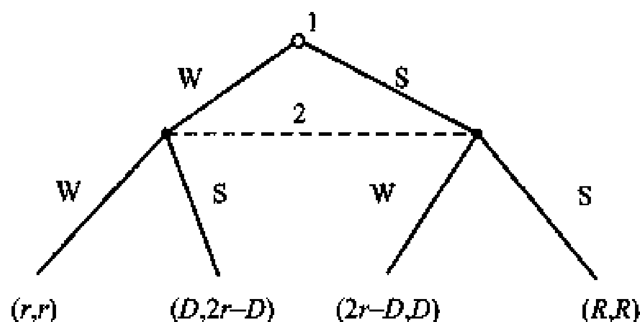


图 4 - 20　基金赎回(2)

现在分析短期的情况。从图 4 - 20 可以看出,这个子博弈存在两个纳什均衡:一个是(赎回,赎回),一个是(不赎回,不赎回)。这说明在基金赎回这个模型中,存在着两个子博弈完美均衡解:① 在短期和长期,两个投资人都赎回,即{(W,W),(W,W)},收益为(r,r)。实际上由于短期都赎回,所以长期并不会发生,但我们仍需要指明参与者(如果走到了长期)的纳什均衡策略。② 在短期,两个投资人都不赎回,而在长期都赎回,即{(S,W),(S,W)},收益为(R,R)。第一个完美均衡实际上就是对基金公司的一次赎回潮,伴随着的就是股市的暴跌和基金公司总资产的加倍缩水。需要指出的是,这个模型虽然简单,但却说明了开放式基金的存在会加大股市的波动,如果开放式基金控制的资金非常庞大,那么这种可能性产生的后果有可能会非常严重。如何防范开放式基金可能的冲击,如何使得投资者的行为长期化,看来是"基金热"下不得不考虑的问题。实际上,长期投资者对于基金公司而言是宝,因而如何识辨短线投机者和长期投资者对于基金公司而言将是十分关键的。

㊥ 习题 4.13　动态性别战。一对情侣决定是在家读书还是去听音乐会。

男首先决定是在家读书还是去听音乐会，如果读书，博弈结束，两人各得 2 的收益；如果去听音乐会，那么男向女发出邀请，并且两人同时决定是去听古典乐还是去听流行乐。如两人分开，那么收益都为 0；如果一起去听古典乐，男得 3 的收益，女得 1；如果一起去听流行乐，男得 1，女得 3。请概括出该博弈的扩展式，并解子博弈完美均衡。

六、国际贸易与关税

各国间的贸易政策和关税协定实际上是一个博弈的结果，中国加入世界贸易组织的谈判过程就清楚地表明了这一点。这里我们考虑一个简单的关税博弈模型。假设有两个国家，每个国家有一个企业，定义为 $i=1$ 和 2。每个国家的政府制定关税税率，各国企业为国内市场和国外市场生产某种产品。如果 $i(=1,2)$ 国的市场总需求为 Q_i，则反需求函数为 $p_i(Q_i)=a-Q_i$，国家 i 中的企业为国内市场生产 h_i，出口 e_i，则 $Q_i=h_i+e_j$，即 i 国的总需求量为国内产量加进口。企业的边际成本为常数 c，没有固定成本，那么企业 i 的成本函数为 $C_i(h_i,e_i)=c(h_i+e_i)$，另外，产品出口时需要负担关税成本：如果政府 j 制定的关税为 t_j，那么企业 i 向 j 国出口 e_i 必须支付的关税就为 t_je_i。首先，两国政府同时选择关税税率 t_1 和 t_2，然后，两国的企业观察到关税，并确定产量 h_i 和 e_i。该博弈的扩展式 $\Gamma=\{N,H,P,u\}$ 为：

(1) 参与者集合：$N=\{$国 1，国 2，企业 1，企业 2$\}$；

(2) 全历史集合：$H=\{(t_1,t_2),[(h_1,e_1),(h_2,e_2)]\}$；

(3) 参与者函数：$P(\varnothing)=\{$国 1，国 2$\}$，$P(t_1,t_2)=\{$企业 1，企业 2$\}$；

(4) 偏好：两国政府的收益函数为社会福利函数，福利由消费者剩余、企业利润和关税组成。根据微观经济学的知识，消费者剩余为消费者获得的总效用减去支出。由于需求曲线是线性的，消费者剩余等于

$$\begin{aligned}\int_0^Q[(a-x)-p]\mathrm{d}x &= \int_0^Q[(a-x)-(a-Q)]\mathrm{d}x \\ &= -\frac{1}{2}Q^2+Q^2=\frac{1}{2}Q^2 \\ &= \frac{1}{2}(h_i+e_j)^2\end{aligned}$$

因而第 i 国政府的福利函数为

$$w_i(t_i,t_j,h_i,e_i,h_j,e_j)=\frac{1}{2}Q_i^2+\pi_i(t_i,t_j,h_i,e_i,h_j,e_j)+t_ie_j$$

企业的收益函数则为利润函数，第 i 国企业的利润函数为

$$\pi_i(t_i,t_j,h_i,e_i,h_j,e_j)=[a-(h_i+e_j)]h_i+[a-(e_i+h_j)]e_i-c(h_i+e_i)-t_je_i$$

因为该博弈是对称的，所以只需考查 i 国和 i 企业即可。

根据逆推法，我们首先考虑在给定两国政府制定的关税税率 t_1 和 t_2 下，两个企业的纳什均衡。对于企业 i，(h_i, e_i) 必须满足

$$\max_{h_i, e_i} \pi_i(t_i, t_j, h_i, e_i, h_j, e_j)$$

由于企业的利润是两个市场利润之和，企业 i 在市场 i 的利润只是 h_i 和 e_j^* 的函数，而在市场 j 的利润却是 h_j^*，e_i 和 t_j 的函数，企业 i 的利润最大化函数可以拆分为两个问题，即在每一个市场分别求最大化。

国内市场 h_i 必须满足

$$\max_{h_i} h_i[a - (h_i + e_j^*) - c]$$

国外市场 e_i 必须满足

$$\max_{e_i} e_i[a - (h_j + e_i) - c] - t_j e_i$$

假设 $e_j^* \leqslant a - c$ 和 $h_j^* \leqslant a - c - t_j$，可得

$$\begin{cases} h_i^* = \dfrac{1}{2}(a - e_j^* - c) & (4-\mathrm{T}-1) \\ e_i^* = \dfrac{1}{2}(a - h_j^* - c - t_j) & (4-\mathrm{T}-2) \end{cases}$$

利用对称性，对于每一个企业 $i = 1, 2$，都必须满足 $(4-\mathrm{T}-1)$ 和 $(4-\mathrm{T}-2)$ 这两个最优反应函数，从而对四个未知数 (h_1, e_1, h_2, e_2) 存在四个方程式。求解非常容易。其解为

$$h_i^* = \frac{a - c + t_i}{3} \quad 和 \quad e_i^* = \frac{a - c - 2t_j}{3} \qquad (4-\mathrm{T}-3)$$

从上述两个解可以看出，如果没有关税，那么两个企业生产的产量等于古诺模型中的产量，即在国内和国外都生产 1/3 的市场容量。但由于关税的存在，使得国内企业的生产多增加了 $t_i/3$ 的产量，即受到了国内关税的保护。相反，在国外则由于关税的保护而受到了损失，产量减少 $2t_j/3$。此外，h_i^* 随着国内关税 t_i 的提高而增加，进口 e_j^* 则随着国内关税的提高而减少，因而关税的提高实际上会增加国家间的锁国性，而不会是外向性，这实际上是一个囚徒困境。

在第一阶段，政府要解出最优的税率以使福利最大化。由于 $(h_1^*, e_1^*, h_2^*, e_2^*)$ 都是税率的函数，所以社会福利是 t_i 和 t_j 的函数。即

$$\max_{t_i} w_i(t_i, t_j)$$

$$\max_{t_i} \frac{(2(a-c) - t_i)^2}{18} + \frac{(a - c + t_i)^2}{9} + \frac{(a - c - 2t_j^*)^2}{9} + \frac{t_i(a - c - 2t_i)}{3}$$

求一阶条件，得

$$t_i^* = \frac{a-c}{3}$$

利用对称性，这一结果对每一个 i 都成立，并不依赖于 t_j，这表明各国政府选择$\frac{a-c}{3}$的税率是严格优策略。把 $t_1 = t_2 = \frac{a-c}{3}$ 带入(4－T－3)式可得

$$h_i^* = \frac{4(a-c)}{9} \quad 和 \quad e_i^* = \frac{a-c}{9}$$

这就得到企业在第二阶段所选择的产量。因而，关税博弈中的子博弈完美均衡解为：第一阶段各国政府选择$\frac{a-c}{3}$的关税，第二阶段企业在国内生产$\frac{4(a-c)}{9}$，在国外生产$\frac{a-c}{9}$，即$\left[\frac{a-c}{3},\frac{a-c}{3},\left(\frac{4(a-c)}{9},\frac{a-c}{9}\right),\left(\frac{4(a-c)}{9},\frac{a-c}{9}\right)\right]$。

实行关税对一国的福利存在什么样的影响呢？不妨假设关税为零，那么各国市场上的商品总量为$\frac{2(a-c)}{3}$，消费者剩余为$\frac{4(a-c)^2}{9}=\frac{32(a-c)^2}{81}$，而在税率为$\frac{a-c}{3}$的情况下，各国的商品总量为$\frac{5(a-c)}{9}$，消费者剩余为$\frac{25(a-c)^2}{81}$，即零关税带来更大的消费者剩余。其次，我们求各国福利之和的最大值，即

$$\max_{t_1,t_2} w_1(t_1,t_2) + w_2(t_1,t_2)$$

利用前面分析的结果，解一阶条件，得

$$-(a-c) = t_i$$

由于税率不可能为负数，所以 $t_1 = t_2 = 0$。即实行零关税，不仅能增加各国的福利，而且能够增加两国的总体福利。由此我们看到，实行关税是一个完美均衡，但对各国和国际社会而言却是无效率的。因而，各国政府就有充分的理由来签订一个相互承诺零关税的协定，这就是为什么 WTO 能够不断发展的根本原因。

(易) 习题 4.14　出口补贴。如果两国政府不仅制定关税税率，而且同时决定出口补贴率 s_i，企业 i 出口 e_i，将能获得 s_ie_i的补贴，那么上述模型会如何变化？

七、工作竞赛

考虑有两个工人，每一个工人 i 的产出取决于其努力程度 e_i和随机扰动项 ε_i，即 $y_i = e_i + \varepsilon_i$。生产的程序如下：第一，两个工人同时选择非负的努力水平 $e_i \geq 0$；第二，随机扰动项 ε_1 和 ε_2 相互独立，服从期望值为0，密度函数为$f(\varepsilon)$ 的

概率分布;第三,工人的产出可以观察到,但工人的努力水平无法观察,从而工人的工资只能决定于个人的产出,而无法直接取决于其努力水平。

假设企业主为了激励工人努力工作而在他们之间开展工作竞赛。产出高的工人即是工作竞赛的优胜者,获得高工资 w_H,而失败者获得低工资 w_L。工人获得工资水平 w 并付出努力程度 e 时的收益为 $u(w,e)=w-g(e)$,其中 $g(e)$ 表示努力工作带来的负效用,是递增的凸函数,即 $g'(e)>0$,且 $g''(0)>0$。企业主的收益为 $y_1+y_2-w_H-w_L$。

定义企业主为参与者 1,他的行动是选择工作竞赛中的工资水平 w_H 和 w_L,两个工人为参与者 2 和参与者 3,他们观察到第一阶段选定的工资水平,然后同时选择努力程度 e_1 和 e_2。最后,参与者各自的收益如前面所给出。由于产出不仅是参与者行动的函数,而且还受到随机扰动因素 ε 的影响,因而参与者的收益为期望值。

该博弈的扩展式 $\Gamma=\{N,H,P,u\}$ 为:

(1) 参与者集合:$N=\{1,2,3\}$;

(2) 全历史集合:$H=\{(w_H,w_L),[e_1,e_2]\}$;

(3) 参与者函数:$P(\varnothing)=1,P(w_H,w_L)=\{2,3\}$;

(4) 偏好:企业主的收益函数为 $U=y_1+y_2-w_H-w_L$,工人的收益函数为 $u(w,e)=w-g(e)$。

运用逆推法。给定企业主选定的工资率 w_H 和 w_L,如果努力水平 (e_1^*,e_2^*) 为第二阶段的纳什均衡,则对每一工人 i 而言,e_i^* 一定使他的期望工资减去努力带来的负效用后的净收益最大化,即

$$\max_{e_i} w_H \text{Prob}\{y_i(e_i)>y_j(e_j)\}+w_L \text{Prob}\{y_i(e_i)\leqslant y_j(e_j)\}-g_i(e_i)$$
$$=(w_H-w_L)\text{Prob}\{y_i(e_i)>y_j(e_j)\}+w_L-g(e_i) \qquad (4-\text{R}-1)$$

其中 $y_i=e_i+\varepsilon_i$。(4-R-1)式的一阶条件为

$$(w_H-w_L)\frac{\text{dProb}\{y_i(e_i)>y_j(e_j^*)\}}{\text{d}e_i}=g'(e_i) \qquad (4-\text{R}-2)$$

(4-R-2)式实际上就是边际收益等于边际成本,边际收益为奖励工资 (w_H-w_L) 与获得奖励工资的概率的乘积(期望值),而边际成本就是边际负效用。

根据贝叶斯法则

$$\text{Prob}\{y_i(e_i)>y_j(e_j)\}=\text{Prob}\{\varepsilon_i>e_j+\varepsilon_j-e_i\}$$
$$=\int_{-\infty}^{+\infty}\text{Prob}\{\varepsilon_i>e_j^*+\varepsilon_j-e_i\mid\varepsilon_j\}f(\varepsilon_j)\text{d}\varepsilon_j$$

$$= \int_{-\infty}^{+\infty} [1 - F(e_j^* + \varepsilon_j - e_i)] f(\varepsilon_j) \mathrm{d}\varepsilon_j$$

于是(4－R－2)式可化为

$$(w_H - w_L) \int_{-\infty}^{+\infty} f(e_j^* + \varepsilon_j - e_i) f(\varepsilon_j) \mathrm{d}\varepsilon_j = g'(e_i)$$

由于我们不清楚具体的函数形式,因而这里要运用到一个技巧:即由于随机扰动项对两个工人的影响的期望值都等于0,因而如果给定工人j的努力程度为e_j,那么工人i的努力程度e_i只要比e_j大一个无穷小正数就能获胜(从期望值来看),因而最优的努力程度是$e_i = e_j$,所以第二阶段的纳什均衡意味着$e_i^* = e_j^* = e^*$。从而上式可化为

$$(w_H - w_L) \int_{-\infty}^{+\infty} f(\varepsilon_j)^2 \mathrm{d}\varepsilon_j = g'(e_i^*) \qquad (4-R-3)$$

由于$g(e)$是凸函数,优胜者获得的奖励越高(即$w_H - w_L$越大),就会激励工人更大的努力,这和我们的直觉是一致的。另一方面,在同样的奖励水平下,随机扰动因素越大,就越不值得努力工作,因为这时工作竞赛的最终结果很大程度上取决于运气而不是努力。不妨设ε服从正态分布,则有

$$\int_{-\infty}^{+\infty} f(\varepsilon_j)^2 \mathrm{d}\varepsilon_j = \frac{1}{2\sigma\sqrt{\pi}}$$

利用反函数,得

$$e^* = g'^{-1}\left(\frac{w_H - w_L}{2\sigma\sqrt{\pi}}\right) \qquad (4-R-4)$$

给定工人的最优选择,企业主如何决定w_H和w_L?根据题意,企业主的期望收益为

$$E[U] = 2e^* - w_H - w_L \text{(随机扰动项的期望值为0)}$$

一阶条件为

$$\frac{\mathrm{d}E[U]}{\mathrm{d}w_H} = \frac{1}{\sigma\sqrt{\pi}} \frac{1}{g''(e^*)} - 1 = 0 \qquad (4-R-5)$$

$$\frac{\mathrm{d}E[U]}{\mathrm{d}w_L} = \frac{-1}{\sigma\sqrt{\pi}} \frac{1}{g''(e^*)} - 1 < 0 \qquad (4-R-6)$$

(4－R－6)意味着w_L越小,企业主的收益越大,但w_L不可能小于0,所以$w_L = 0$。这时,e^*就是w_H的函数,只要给出具体的函数$g(e^*)$,就能解出w_H和e^*。因而该博弈的子博弈完美均衡为:[((4－R－5),0),(4－R－4),(4－R－4)]。

如果我们对该博弈稍作修改,例如允许工人退出,那么只要有一个工人退出,竞赛将无法进行,企业主得0,工人得到保留收益R。如果工人同意参与工作竞赛而不是另谋高就,他们将根据(4－R－4)来选择努力水平。因为在对称

的纳什均衡中($e_i^*=e_j^*$),每个工人在竞赛中获胜的概率就为 $\text{Prob}\{y_i>y_j\}=1/2$,因而企业主要使工人有积极性参加工作竞赛,则他选择的工资率必须满足下式

$$\frac{1}{2}w_H+\frac{1}{2}w_L-g(e^*)\geqslant R \qquad (4-R-7)$$

(4-R-7)式被称为参与约束,它表示要吸引工人参与博弈必须满足的条件。假如 R 足够低,以至于工人愿意参加竞赛,则企业主在(4-R-7)式的约束下,选择使自己期望收益 $2e^*-w_H-w_L$ 最大的工资水平。由于在最优条件下,(4-R-7)中的等号成立,所以

$$w_L=2R+2g(e^*)-w_H \qquad (4-R-8)$$

则企业主的期望利润就成为 $2e^*-2R-2g(e^*)$,于是企业主要考虑的问题就是使 $e^*-g(e^*)$ 最大化,这时他选择的工资水平应激励工人选择 e^* 以满足这一条件。从而最优条件下的努力水平满足一阶条件 $g'(e^*)=1$,将其代人(4-R-3)则意味着最优激励 w_H-w_L 满足

$$(w_H-w_L)\int_{-\infty}^{+\infty}f(\varepsilon_j)^2\mathrm{d}\varepsilon_j=1$$

即 $w_H-w_L=2\sigma\sqrt{\pi}$,与(4-R-8)一起,如果知道 $g(e^*)$ 的具体形式,就可解出 w_H^* 和 w_L^* 的值。因而该博弈的完美均衡解为$[(w_H^*,w_L^*),e_1^*,e_2^*]$,其中 $g'(e^*)=1$。

㊀ 习题 4.15 如果给定 $g(e)=e^2$,$R=0$,随机扰动服从正态分布,那么工作竞赛的子博弈完美均衡是什么?

八、退出衰落行业

设一个行业的市场需求函数为 $P_t(Q)$,$t=1,2,3,\cdots$,并且 $P_t(Q)<P_{t+1}(Q)$,即该行业是一个衰落行业。在该行业存在两个垄断企业,不失一般性,q_1 代表第1个企业的产能,q_2 代表第2个企业的产能,其中 $q_1>q_2$,垄断企业 i 要么不生产,要么选择产量 q_i,生产的总成本为 cq_i,垄断企业不生产就意味着退出该行业,退出之后不可能再进入。任何一个时期,市场上的商品总量为两个企业的产量之和,即 $Q=q_1+q_2$。随着时间的推移,市场为逐渐萎缩,因而企业会陆续退出,那么企业在什么时候退出?谁先退出?为了回答这些问题,我们先将其概括为扩展式,然后再寻找其于博弈完美均衡。

该博弈的扩展式 $\Gamma=\{N,H,P,u\}$ 为:

(1) 参与者集合:$N=\{1,2\}$。

(2) 全历史集合:$H=\{(X^1,X^2,\cdots,X^t)\}$。令退出 = E,继续 = S。全历史可

能出现三种情况：如果 $X^t=(E,E)$，那么 $X^r=(S,S)$，$t>r\geqslant 1$。如果 $X^t=E$，那么当 $t>s>r$ 时，$X^s=S$；当 $s=r$ 时，$X^s=(S,E)$ 或 (E,S)；当 $r>s\geqslant 1$ 时，$X^s=(S,S)$。如果 $X^t=E$，$X^{t-1}=(S,E)$ 或 (E,S)，那么 $X^r=(S,S)$，$t-1>r\geqslant 1$。

(3) 参与者函数：$P(\varnothing)=\{1,2\}$，对于任一历史 h，$P(h)=\{1\}$ 或 $\{2\}$ 或 $\{1,2\}$。

(4) 偏好：每个垄断企业都追求利润最大化。由于单位成本为常数 c，因而利润的正负实际上就取决于 $p(Q)-c$，如果两个企业都生产，那么利润为正就要求 $p(q_1+q_2)>c$，如果一个企业生产，那么利润为正就要求 $p(q_i)>c$。

我们上面提出的问题实际上等价于该博弈的子博弈完美均衡，而要寻找该博弈的子博弈完美均衡，就需要运用逆推法，虽然不知道该博弈什么时候结束，但由于该行业在不断萎缩，因而仍可以运用逆推法来求解。不妨设 t_0 为两个企业都生产，并且利润都为正的最后一个时期；t_1 则为只有垄断企业 1 生产，且利润为正的最后一个时期；同理，t_2 则为只有垄断企业 2 生产，且利润为正的最后一个时期。那么根据题意，可知 $t_0\leqslant t_1\leqslant t_2$，如图 4－21 所示。

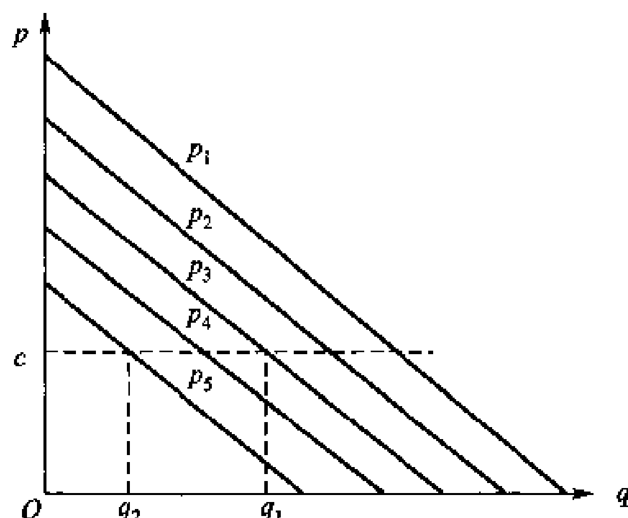


图 4－21　衰落行业的需求函数

在图 4－21 中，$t_0=1$，$t_1=3$，$t_2=5$。不失一般性，我们只考虑 $t_0<t_1<t_2$ 的情况。

(1) 当 $t>t_2$ 时，显然，假设还有企业没有退出该行业，那么该企业必然选择退出该行业。

(2) 当 $t_1>t\geqslant t_2$ 时，显然，即使只有企业 1 留在该行业，它也是亏本的，因而无论如何它的最优选择都是退出该行业。在这种情况下，企业 2 最优选择就是留在该行业。

(3) 当 $t=t_1$ 时，如果两个企业都留在该行业，那么必然亏损，因而市场只能

容许一个企业存在。如果企业2选择留在该行业，那么企业1的最优选择只能是退出该行业。当$t=t_1-1$时，同理可得企业2留（能得到t_1到t_2的利润之和），而企业1退（避免亏损）。因而，在$t(t_0>t\geq t_1)$期，企业1都会选择退，企业2都会选择留。此外，如果企业1在t_1选择留在该行业（希望企业2退出），那么企业2是否退出该行业呢？如果企业2选择退出，那么它将失去t_1+1到t_2的利润之和，因而是否退出关键就在于企业2在t_1留下产生的亏损与t_1+1到t_2的利润之和谁大谁小，这又需要考虑三种情况。如果利润之和大于亏损，那么企业2不论企业1如何做，都会留下，因而企业1只能退出。如果利润之和小于亏损，那么企业2（在企业1留下的情况下）只能退出。如果亏损与利润之和相等，那么对于企业2而言，留和退都一样，如果留，那么企业1只能退；反之，如果企业2退，企业1就会留。我们将其逆推回t_0+1期，可知：① 企业2的t_1+1到t_2的利润之和大于t_1期的亏损，唯一的均衡策略为在$t(t_0>t\geq t_1)$期（企业1退，企业2留）。② 如果企业2的t_1+1到t_2的利润之和小于t_1期的亏损，那么存在两种均衡策略：在t期（企业1留，企业2退）或（企业1退，企业2留）。③ 如果企业2的t_1+1到t_2的利润之和等于t_1期的亏损，同样存在两种可能：一种是在每个t期（企业1退，企业2留），一种是在每个t期（企业1留，企业2退）。总结一下：无论什么情况，在每一个$t(t_0>t\geq t_1)$期，（企业1都会选择退，企业2都会选择留）是一个均衡结果；如果企业2的t_1+1到t_2的利润之和$\leq t_1$期的亏损，那么还可能存在另一个均衡结果：在每一个$t(t_0>t\geq t_1)$期（企业1都会选择留，企业2都会选择退）。

（4）当$t_0\geq t\geq 1$时，显然，两个企业都会选择留在该行业。

综上所述，该博弈的子博弈完美均衡如下：无论什么情况，（在$1\sim t_0$期，企业1选择留，在$t>t_0$的每一期，企业1都选择退；在$1\sim t_2$期，企业2选择留，在$t>t_2$的每一期，企业2都选择退），如果企业2的t_1+1到t_2的利润之和$\leq t_1$期的亏损（两个企业都存在），那么还可能存在另一个子博弈完美均衡（在$1\sim t_1$期，企业1选择留，在$t>t_1$的每一期，企业1都选择退；在$1\sim t_0$期，企业2选择留，在$t>t_0$的每一期，企业2都选择退）。现在我们就可以回答最早提出的问题，企业最早退出的时间发生在t_0期，至于哪个企业最早退出则需要根据情况来定，如果企业1退出，那么企业2将在t_2时退出。此外，如果企业2的t_1+1到t_2的利润之和$\leq t_1$期的亏损，那么除了上述提到的退出时间表，还可能存在着另外一个时间表，即企业2在t_0退出，而企业1在t_1退出。总之，在t_0期之后该行业只可能存在一个企业，该行业的消失则可能发生在t_1或是t_2期。

实际上，在t_0期之后，哪个企业留在该行业还取决于各企业的财务状况，例如，如果企业2在每一期都要求利润不能为负，那么企业1就会在t_0期之后仍然

留在该行业。如果企业 2 的负债有一个限额，以至于企业 2 在 t_0期之后只能坚持较短时期（不妨设为 L），这时如果企业 1 在 $t_0+1 \sim t_0+L$ 时期的亏损（企业 2 仍留在该行业）能够被 $t_0+L+1 \sim t_1$时期的利润所弥补（企业 2 退出该行业），企业 1 就会坚持留在该行业，面对这种情况，企业 2 的最优策略只能是在t_0+1期退出。如果我们定义一个 k，使得企业 1 在 $t_0+1 \sim t_0+k$ 时期的亏损等于 $t_0+k+1 \sim t_1$时期的利润之和，那么这个 k 就可以理解为企业 2 的负债限额，只要 $L<k$，那么企业 1 就值得继续留在该行业。

(易) 习题 4.16　请你考虑更为一般的情况，即 $t_0 \leqslant t_1 \leqslant t_2$的情景。

(易) 习题 4.17　如果 $c=10, q_1=20, q_2=10$，需求函数 $p(Q)=100-t-Q>0$，否则为 0。找出所有的子博弈完美均衡，如果企业 2 的负债有一个限额，那么求出 k。

关键词

完全信息动态博弈　完美信息动态博弈　全历史　空历史　子历史　真子历史　参与者函数　动态博弈的扩展式 $\Gamma=\{N,H,P,u\}$　策略和结果　结果函数　不可置信威胁　动态不一致　扩展式博弈的纳什均衡　扩展式博弈的策略式　动态博弈的子博弈 $\Gamma(h)$　子博弈完美均衡　逆推法　扩展式博弈的有限性

内容要点

1. 完全信息动态博弈与完美信息动态博弈的区别就在于前者包含着参与者同时行动这样的情况。在完全信息下，所有的历史都能被观察到并被记忆住。

2. 如果具有了四个要素：参与者集合、全历史集合、参与者函数和参与者的收益函数 u，就可以完整地描述一个完全信息动态博弈。所以说，完全信息动态博弈也可以根据是否具备这四个要素来加以判断。

3. 策略是参与者的一种“万全之策”，它不仅需要指明在可能的情况下，参与者如何做，而且需要指明在“万一”的情况下参与者如何做。策略不再是行动。结果是指在相应的策略下“实际”发生的历史情况，显然它与策略不同，但却存在着对应关系，这种对应关系可用结果函数来表达。

4. 在动态博弈中，运用纳什均衡的概念会产生一些问题，这主要是纳什均衡的要求不够严格造成的。动态不一致的现象在纳什均衡中有可能出现，而不可置信威胁是动态不一致的一种主要表现形式，为了克服动态不一致，引入了比

纳什均衡要求更强的概念——子博弈完美均衡。

5. 子博弈完美均衡要求一个策略组合必须是所有子博弈(包括原博弈)的纳什均衡,而动态博弈中的纳什均衡仅要求在一个子博弈(原博弈)中它是纳什均衡。它与纳什均衡的关系在于:子博弈完美均衡一定是纳什均衡,但反之不成立。

6. 在有限动态博弈中,逆推法与子博弈完美均衡完全等价,并成为求解动态博弈中的子博弈完美均衡的主要方法(有时甚至是唯一的方法)。

进一步阅读

我们这里所采用的扩展式表达强调历史而忽略节点,这种表达法可参阅 Osborne(2004)和 Osborne and Rubinstein(1994),另一种较流行的扩展式表达则是根据博弈树中的节点来进行定义的,对其的介绍可参阅 Kreps(1990)和 Gibbons(1992)。

第五章

重复博弈

如果囚徒困境这个博弈不是进行一次，而是进行（无穷）多次，那么（沉默，沉默）这样的策略组合会不会作为一种均衡结果在博弈中出现呢？换言之，如果将囚徒困境稍加改造，比如，沉默换成诚信，而坦白改成欺骗，那么（诚信，诚信）这样的非均衡结果会不会作为一种理想状态在社会中长期存在？将（欺骗，欺骗）的均衡结果演变为（诚信，诚信）这样的理想状态，无疑是一种帕累托改进，但实现这种改进的机制又是什么？对上述问题的回答，就涉及本章的内容——重复博弈。重复博弈不仅在博弈论中有着重要的理论价值，而且是理解人类社会合作的基础，或者说为亚当·斯密“看不见的手”都提供了一种深刻而统一的解说。

这一章将从重复博弈的定义入手，通过对合作解如何产生这一问题的追寻，最终给出无名氏定理，最后则通过经典例子来帮读者加深对理论的理解和运用。

第一节　重复博弈的定义及扩展式

顾名思义，重复博弈就是相同的博弈 G 在长期中不断地重复进行，博弈的任何阶段实际上都是一个完全信息静态博弈，为了方便，把重复博弈中具有相同片段的博弈 G 称为阶段博弈。重复博弈与通常的完全信息动态博弈的一个显著差异就是在博弈的任意一个阶段，都会产生相应的收益。例如，一个重复无限次的囚徒困境，博弈的每一个阶段都会产生相应的收益，从而形成一个收益流。显然，当前的收益与未来的收益对于博弈参与者的效用应当是不同的，其中存在着一个贴现的问题。

一、贴现因子与偏好

明天的一元钱和今天的一元钱价值是不一样的，最简单的理由是今天的一元钱如果存入银行那么在明天会变成 $1+r$，所以明天的一元钱只相当于今天的

$1/(1+r)$元钱，$1/(1+r)$实际上就是经济学中的贴现率。如果未来存在收益序列 $R_1,R_2,R_3,\cdots$，那么这个未来收益流的贴现值之和就为

$$V = R_1 + \delta R_2 + \delta^2 R_3 + \cdots = \sum_{t=1}^{\infty} \delta^{t-1} R_t \qquad (5-1)$$

其中 $\delta \in (0,1)$ 称为贴现因子(discount factor)。严格讲，贴现因子并不等于贴现率，但贴现因子与贴现率一定是同方向变动的。例如，考虑一个特殊的重复博弈，其结束之前重复进行的次数是随机的，即在博弈的每一阶段完成之后，都要通过抛若干枚(加权的)硬币的方式来决定博弈是否结束，如果硬币朝上那么博弈结束(即概率为 p)，如果是其他情况，那么博弈继续(即概率为$1-p$)。如果下一阶段能得到的收益为 R_1，那么在当前阶段硬币未抛之前的价值(即贴现后的期望值)为$(1-p)R_1/(1+r)$；如果下两阶段能得到的收益为 R_2，在当前阶段硬币未抛之前的价值为$(1-p)^2R_2/(1+r)^2$；下三阶段、四阶段等的收益照此类推。令 $\delta=(1-p)/(1+r)$，则贴现因子既包含了货币的时间价值(贴现率 $1/(1+r)$)，又包含了博弈结束的可能性$(1-p)$。

有了贴现因子的概念，就可以比较无穷重复博弈中的不同收益值，从而对不同的策略进行优劣判断。例如考虑无限重复的囚徒困境，在每一阶段参与者可以有不同的行动，得到的收益也会不一样，在没有贴现因子的情况下，其在无限重复博弈中的值之和将等于无穷大，分析起来十分不便，甚至无法对参与者的不同策略进行比较。但引入贴现因于 δ 后，未来收益流的贴现值必然不是无穷数，即无穷级数(5-1)是收敛的(只要每一期的收益 R_t 有限)。例如，如果 t 期的收益为 R_t，贴现因子为 $1>\delta>0$，那么收益流的贴现值为

$$R_1 + \delta R_2 + \delta^2 R_3 + \cdots = \sum_{t=1}^{\infty} \delta^{t-1} R_t < \lim_{n\to\infty} R_{max}(1+\delta+\delta^2+\cdots+\delta^n) = \frac{R_{max}}{1-\delta}$$

其中 $R_{max}=\max\{R_1,R_2,R_3,\cdots\}$，即 R_{max}为收益流中的最大值。同理

$$R_1 + \delta R_2 + \delta^2 R_3 + \cdots = \sum_{t=1}^{\infty} \delta^{t-1} R_t > \lim_{n\to\infty} R_{min}(1+\delta+\delta^2+\cdots+\delta^n) = \frac{R_{min}}{1-\delta}$$

其中 $R_{min}=\{R_1,R_2,R_3,\cdots\}$，即 R_{min}为收益流中的最小值。就这意味着，存在一个 R 使得

$$R_1 + \delta R_2 + \delta^2 R_3 + \cdots = \sum_{t=1}^{\infty} \delta^{t-1} R_t = \lim_{n\to\infty} R(1+\delta+\delta^2+\cdots+\delta^n) = \frac{R}{1-\delta}$$

R 就被称为收益流$(R_1,R_2,R_3,\cdots)$的贴现平均收益值。

定义 5.1 设贴现因子为 δ，收益流$(R_1,R_2,R_3,\cdots)$的**贴现平均收益值**为

$$R = (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} R_t$$

由于平均收益值等于贴现值之和 V 的 $(1-\delta)$ 倍，使贴现平均收益值最大化就等同于使贴现值之和最大化。使用平均收益的另一个优点，就是可以利用它直接和阶段博弈中的收益进行比较，从而更容易知道哪一个策略要优。

在前面的分析中，我们都是用收益函数（效用函数）来描述博弈参与者的偏好。在确定性下，表达相同偏好的收益函数并不唯一，而是满足单调变换性，即只要 f 是一个单调递增函数，那么 $f(u(\cdot))$ 与 $u(\cdot)$ 表示同一个偏好。在重复博弈中，整个博弈的收益函数为

$$v = u(s^1) + \delta u(s^2) + \delta^2 u(s^3) + \cdots = \sum_{t=1}^{\infty} \delta^{t-1} u(s^t) \qquad (5-2)$$

它实际上为阶段博弈 G 的收益函数 $u(s)$ 的一个贴现和，我们把 $u(s)$ 也称为**伯努利收益函数**，因为它也像 v－N－M 偏好一样，要求 $u(s)$ 必须满足线形变换，即只有当 $f=a+bu(s)$，$b>0$ 时，f 和 u 才表示相同的重复博弈偏好。这一点与贴现因子包含随机因素相一致，因为这时的 v 实际上是预期收益函数。因而，重复博弈与通常的完全信息动态博弈的第二个不同点，就是收益函数 $u(\cdot)$ 为伯努利收益函数，而不是普通的收益函数。

(易) 习题 5.1　请你证明为什么重复博弈中的收益 v 必须为 v－M－N 收益函数。（提示：用 $(1-\delta)$ 乘 $(5-2)$ 式）

二、重复博弈的定义及扩展式

有了前面理论的铺垫，现在我们就开门见山地给出重复博弈的定义。

定义 5.2　对于策略式博弈 $G=\{N,S,u\}$，其中 $N=\{1,2,\cdots,n\}$ 为参与者集合，$S=\{S_1,\cdots,S_n\}$ 为所有参与者的策略空间（策略实际上就是行动），$u=\{u_1,\cdots,u_n\}$ 为所有参与者的收益函数，如果 G 在时间中（或程序上）不断重复，并且在下一次博弈 G 开始前，所有以前博弈的历史都被观察到，那么它构成的动态博弈就称之为**重复博弈**，G 就为重复博弈中的**阶段博弈**。如果 G 重复进行 T 次，那么 $G(T)$ 就表示重复进行 T 次的有限重复博弈。如果 $T=\infty$，那么 $G(T)$ 就表示无限重复博弈。重复博弈 $G(T)$ 中参与者 i 的偏好用收益函数 v_i 表示，即

$$v_i = u_i(s^1) + \delta u_i(s^2) + \cdots + \delta^{T-1} u_i(s^T) = \frac{R_i}{1-\delta} \qquad (5-3)$$

其中 $u(s^t)$ 为伯努利收益函数，s^t 为重复博弈 t 阶段的行动组合 $(T \geqslant t \geqslant 1)$，$\delta$ 为贴现因子，R_i 为参与者 i 的贴现平均收益值，等于 $(1-\delta)\sum_{t=1}^{T}\delta^{t-1}u_i(s^t)$。

实际上，对于重复博弈中的阶段博弈 G，其不仅仅为完全信息静态博弈，也可以是完全信息动态博弈；不仅可以为完全信息博弈，也可以为非完全信息博

弈。相应的重复博弈 $G(T)$ 的扩展式定义如下：

定义 5.3　重复博弈的扩展式为 $\Gamma(T)=\{N,H,P,v\}$，其中 N 表示参与者集合；H 为全历史集合，即 G 中的行动组合序列的集合 $\{(s^1,s^2,\cdots,s^T)\}$，有时也把一个全历史称为一个**结果路径**；P 为每一个子历史 h 下的参与者函数；v 则为参与者的 v－N－M 收益函数，它满足(5－3)式。

为了叙述简便，$G(T)$ 与 $\Gamma(T)$ 通常互换使用，都表示重复博弈。

第二节　合作产生的原原

为了更为形象，我们引入一个重复信用困境博弈，其阶段博弈 G 的博弈矩阵如图 5－1 所示。对重复信用困境博弈的分析，将贯穿本章的始终。

		商人 2	
		诚信	欺骗
商人 1	诚信	4,4	0,5
	欺骗	5,0	1,1

图 5－1　信用困境博弈矩阵(1)

一、问题的提出

在什么样的情况下，(诚信，诚信)这样的结果有可能在信用困境中出现？这里，不妨考虑一个简单的两阶段重复信用困境博弈。由于该博弈只进行两次，为了简单，不妨假设贴现因子 $\delta=1$。在重复两次信用博弈中的第二阶段，如果两个商人都讲诚信的话，那么他们的福利都能改善(相对于纳什均衡(欺骗，欺骗))，各得 4 的收益而不是 1 的收益；但如果一方讲诚信，那么另一方就会存在欺骗的动机，因为欺骗能获得更大的利益 5。因而在第二阶段，信用困境的纳什均衡为(欺骗，欺骗)，而不管第一阶段的选择如何。逆推到第一阶段，将第二阶段的纳什均衡收益代入，则得到新的博弈矩阵，如图 5－2 所示。

		商人 2	
		诚信	欺骗
商人 1	诚信	5,5	1,6
	欺骗	6,1	2,2

图 5－2　信用困境博弈矩阵(2)

由于第二阶段的纳什均衡收益为1,所以图5-2中各策略组合对应的收益都加上了1。在第一阶段,唯一的纳什均衡仍是(欺骗,欺骗)。从而,两阶段信用困境唯一的子博弈完美均衡为第一阶段的(欺骗,欺骗)和随后第二阶段的(欺骗,欺骗)。利己利人的合作解——(诚信,诚信)不可能在重复两次的信用困境中出现。那么,如果我们假设信用博弈重复100次,200次,10 000次,上述两阶段分析的结果会不会改变呢?

如果阶段博弈 G 存在唯一纳什均衡,那么 $G(T)$ 的子博弈完美均衡不过是纳什均衡重复T次,根本的原因是从逆推法入手,无论前面的历史如何,最后一个子博弈 $G(1)$ 的纳什均衡唯一,因而 $G(T)$ 的完美均衡不过是 G 的纳什均衡重复 T 次,这就有了命题5.1。

命题5.1 如果阶段博弈 G 有唯一的纳什均衡,则对任意有限的 T,重复博弈 $G(T)$ 有唯一的子博弈完美均衡,即 G 的纳什均衡结果在每一个阶段重复进行。

利用命题5.1可知,无论信用困境重复多少次,只要不是无穷的,那么唯一的均衡结果只能是每一阶段都为(欺骗,欺骗),因而人类社会所谓的合作根本就不可能产生,人与人之间的诚信只能是一种奢望。然而,现实并非如此,虽然人与人之间存在着利益冲突,但也确实存在着合作的行动和结果,否则人类社会早就终结了。这种合作行动不仅在人类社会存在,实际上在动物的行为中同样存在。莎士比亚笔下的夏洛克虽然歹毒,但不可否认他是诚信的。那么如何解释这种合作行动呢?

二、多重均衡的有限重复博弈

为了在理论上容纳合作解,一个重要的方法就是在重复博弈中引入多重均衡。现在,我们仍回到两阶段信用博弈,但对信用博弈进行修改,即阶段博弈 G 存在多个纳什均衡的情况,如图5-3所示,不妨称为信用困境1。信用困境1与原信用困境不同的地方是人为构造了一个新的纳什均衡——(中,中),因而这个阶段博弈存在两个纳什均衡——(欺骗,欺骗)和(中,中)。

		商人2		
		诚信	欺骗	中
	诚信	4,4	0,5	0,0
商人1	欺骗	5,0	1,1	0,0
	中	0,0	0,0	3,3

图5-3 信用困境1(1)

我们仍然规定信用困境1重复进行两次。与原信用困境一个重要的不同是合作解有可能成为子博弈完美均衡解在重复博弈中出现。现在我们就证明在重复两次的信用博弈1中存在一个子博弈完美均衡解,其中每一个商人的策略为:在第一阶段选择诚信;在第二阶段,如果对方在第一阶段选择了诚信,那么选择中;如果不是,那么选择欺骗。

首先从信用困境1的第二阶段开始。在第二阶段,(欺骗,欺骗)和(中,中)都是阶段博弈 G 的纳什均衡,因而"策略组合"[(如果对手诚信,选择中;如果非诚信,选择欺骗),(如果对手诚信,选择中;如果非诚信,选择欺骗)]是子博弈 $G(1)$ 的一个纳什均衡。把这一均衡策略下对应的收益代人第一阶段,就得到图5-4。图5-4中的收益是这样得到的,第一阶段的策略组合(诚信,诚信)对应的收益要各加上3(因为在第一阶段选择了诚信,得到了奖励),而其余的8个策略组合各加上1(因为选择了非诚信,受到惩罚)。

		商人2		
		诚信	欺骗	中
	诚信	7,7	1,6	1,1
商人1	欺骗	6,1	2,2	1,1
	中	1,1	1,1	4,4

图5-4　信用困境1(2)

在图5-4中显然存在着三个纯纳什均衡:(欺骗,欺骗)、(中,中)和(诚信,诚信)。根据逆推法,这3个纯纳什均衡都是信用困境1重复博弈的子博弈完美均衡解。图5-4中的纳什均衡(欺骗,欺骗)对应着重复博弈的完美均衡结果[(欺骗,欺骗),(欺骗,欺骗)],而纳什均衡(中,中)则对应着重复博弈的完美均衡结果[(中,中),(欺骗,欺骗)]。重复两次信用困境1中的这两个完美均衡都简单地由两个阶段博弈的纳什均衡组成,但第三个纳什均衡结果却由一个非纳什均衡(第一阶段)和一个纳什均衡(第二阶段)组成。它对应着的子博弈完美均衡结果为[(诚信,诚信),(中,中)],与前两个完美均衡本质不同的地方是合作解(诚信,诚信)在第一阶段中出现了。实际上,如果 G 是一个有着多重纳什均衡的完全信息静态博弈,则重复博弈 $G(T)$ 就可能存在子博弈完美均衡解,其中对每一个 $t<T$,第 t 阶段的结果都不是 G 的纳什均衡。这就从一个方面揭示了为什么合作解能够在重复博弈中出现,因为合作得到奖励,不合作受到惩罚,因而每一方都会加以遵守。在人类社会中,这种奖励与惩罚机制可以通过国家意志——法律的形式得到实现,从而为人类的合作提供一种制度上的保障。所谓的契约社会,就是通过法律、道德的形式来实现这种诚信(遵守

契约规定）。

但是这种多重纳什均衡的处理仍然存在着令人不满意的地方，因为多重均衡意味着合作解并不一定产生，它的出现需要太多的条件（例如，面对自私的“恶人”，道德约束是软的，而法律约束虽然强硬，但却是事后的，甚至法律本身可能存在漏洞，或者打官司搜集证据的成本高昂，这些都可能导致合作解无法出现），而且合作解对于干扰过于敏感（合作双方很容易产生不信任），相当不稳定，这表明子博弈完美均衡对可信性的要求并不严格。例如，在推导子博弈完美均衡{[诚信,(中,欺骗)];[诚信,(中,欺骗)]}时，我们假定如果第一阶段的结果是（诚信,诚信），则参与者双方都预期（中,中）将是第2阶段的解，如果第一阶段出现了任何其他8种结果之一，第二阶段的结果就会是（欺骗,欺骗）。但是，在第二阶段，（中,中）是一个具有优势的纳什均衡，3的收益显然要大于选择（欺骗,欺骗）时只能得1的收益。因而无论第一阶段情况如何，第二阶段选择中是理性的，过去的并不计人成本。一个很自然的想法是，参与双方可能会重新谈判。每一个参与者可能会理性地认为过去的反正已经过去了，在余下的阶段博弈中就会选择双方都偏好的均衡行动（中,中）。但如果最后一个阶段的结果都是（中,中）的话，那又有什么理由去保证有人在第一阶段去选择（诚信,诚信）呢？{[诚信,(中,欺骗)];[诚信,(中,欺骗)]}就不再是一个“合理的”子博弈完美均衡。

为了克服上述问题，一个思路就是拒绝谈判，为什么拒绝谈判，因为拒绝谈判比接受谈判有更好的利益。不妨考虑如图5-5所示的信用困境2博弈。

		商人2 诚信	欺骗	中	左	右
商人1	诚信	4,4	0,5	0,0	0,0	0,0
	欺骗	5,0	1,1	0,0	0,0	0,0
	中	0,0	0,0	3,3	0,0	0,0
	左	0,0	0,0	0,0	$4,\frac{1}{2}$	0,0
	右	0,0	0,0	0,0	0,0	$\frac{1}{2},4$

图5-5 信用困境2

信用困境2在图5-3的基础上加上了策略左和右，从而阶段博弈G有了四个纯策略纳什均衡：（欺骗,欺骗），（中,中），（左,左）和（右,右）。设想信用困境2博弈重复进行两次，且在第二阶段开始前可以观测到第一阶段的结果。进一步假设商人1（或2）的策略为：第一阶段选择诚信；在第二阶段，如果对方在

第1阶段选择诚信,那么选择中;如果对方在第一阶段没有选择诚信,那么第二阶段选择左(或右)。那么((诚信,诚信),(中,中))就是重复信用博弈2的子博弈完美均衡结果,因为先选诚信,接着选中,每个参与者都可得到4+3的收益,但在第一阶段偏离这一选择而选欺骗,却只能得 $5+\frac{1}{2}$(选择其他行动收益更低)。更为重要的是,前一例子中遇到的困难在这里并没有出现。对于信用困境1而言,对一个参与者在第一阶段不守信用的惩罚,只能通过第二阶段选择一个较劣的纳什均衡来对欺骗者进行惩罚,但在惩罚欺骗者的同时,也惩罚了诚信者,因而在信用困境1中就存在着重新进行谈判的动机。而在信用困境2中,这一点得到了克服,在惩罚欺骗者的同时,对诚信者进行奖励,欺骗者得 $\frac{1}{2}$,而诚信者得 $4>3$,因而诚信者没有动机去进行重新谈判。因而,对于商人1(或2)而言,(第一阶段选择诚信,第二阶段如果对方也诚信,那么选择中;如果对方欺骗,那么选择左(或右))就是子博弈完美均衡,(诚信,诚信)就有可能作为均衡结果出现。

从上述分析可以看出,信用困境1和信用困境2人为设计的痕迹过于明显,毕竟现实生活不是随意设计的游戏。另外一个冲突的地方是,在现实生活中有些博弈只有唯一的纳什均衡,甚至没有纳什均衡,但合作解同样出现,而且具有相当的稳定性。这就引出了无限重复博弈。

㊐ 习题5.2　在有限重复信用困境中,如果参与者选择了某个 x 行动,在下阶段如果其改变行动的话,其将承担 c 的成本,那么 c 应当多大才能保证(诚信,诚信)这一结果有可能在有限重复博弈中出现?你能给出这样的一个策略吗?

三、无限重复信用博弈

合作解要在有限重复博弈中出现要求阶段博弈 G 必须存在多重纳什均衡,但在无限重复博弈中这一条件并不是必需的:即使阶段博弈 G 只存在唯一纳什均衡(甚至不存在纯纳什均衡),无限重复博弈中也存在子博弈完美均衡解,其中没有任何一个阶段结果是 G 的纳什均衡。这显然与命题5.1相对立,根本的原因就在于如果博弈是无限重复的,当 δ 充分接近1,即人们有足够的耐心,那么考虑长远利益就要好于短视。

我们仍然考虑信用困境博弈(见图5-1),但它无限次地进行下去,并且在每 t 期开始之前,$t-1$ 次阶段博弈的结果都已被观察到。其中每一参与者的贴现因子都为 δ,且每一参与者在重复博弈中得到的收益等于各自在所有阶段博弈中所得收益的现值。我们将证明尽管阶段博弈中唯一的纳什均衡是(欺骗,

欺骗），但在无限重复信用博弈中子博弈完美均衡却有可能导致这样的结果：每一阶段都是相互合作——（诚信，诚信）。其思路在于：如果当前参与者选择合作，那么将来他将得到好报（选择高收益的均衡结果）；如果欺骗，将来他将得到报应（选择低收益的均衡结果）。“两害相权取其轻，两利相权取其重”，从而使得合作在信用困境中出现。

每一阶段都重复纳什均衡（欺骗，欺骗）显然是无限重复信用困境的子博弈完美均衡，问题是还存在着其他“更好”的子博弈完美均衡吗？要找出无限重复信用困境博弈中的所有完美均衡是一个不可能完成的任务，我们不妨从现实生活中抽象出一些认为“合理”的策略，然后证明它们是否为子博弈完美均衡。这里我们考虑三个常见的策略：触发策略（trigger strategy），有限惩罚策略（limited punishment stategy）和一报还一报策略（tit for tat strategy）。

（一）触发策略

所谓触发策略是指：如果没有人选择欺骗，诚信将一直进行下去；一旦有人选择欺骗，就会触发其后所有阶段都不再相互诚信（合作）。我们可把参与者 i 的这一策略正式表述为：

在第一阶段选择诚信，且在第 $t>1$ 阶段，如果所有前面 $t-1$ 阶段的结果都是（诚信，诚信），则选择诚信，否则选择欺骗。

图 5－6 形象地说明了触发策略，（诚信，诚信）表示以前的所有历史都是诚信，现阶段仍然选择诚信，如果发生过欺骗将永远选择欺骗。



图 5－6 触发策略

如果参与双方都采取这种触发策略，则无限重复信用博弈的均衡结果就将是每一阶段都选择（诚信，诚信）。我们首先论证如果 δ 充分接近 1，则采取这种策略，对参与双方而言都是最优的，因而为纳什均衡，然后证明这一纳什均衡是子博弈完美均衡。

要证明采取上述触发策略是无限重复信用困境的纳什均衡，等同于证明如果给定参与者 i 的策略为触发策略，那么参与者 j 的最优反应也是触发策略，即触发策略是彼此策略的最优反应。假定参与者 i 已采取触发策略，并证明在 δ 足够接近 1 的条件下，参与者 j 的最优反应同样是触发策略。

第一，假如某阶段的结果偏离了（诚信，诚信），参与者 i 将在其后永远选择欺骗，在这种情况下，参与者 j 的最优反应同样是在其后永远选择欺骗（选择诚信将得 $0<1$）。这表明（欺骗，欺骗）是无限重复信用博弈中参与者双方的最优

反应。

第二，余下的就是计算参与者 j 在第一阶段的最优反应，以及 t 期前的结果都是(诚信，诚信)时，t 期的最优反应。选择欺骗将会使当期得到5的收益，但却会触发参与者 i 的永远不合作(在对方欺骗的情况下，参与者 j 的最优策略只能是欺骗)，于是 t 期后的收益都将成为1，因而欺骗导致的收益流的现值为

$$V = 5 + \delta \times 1 + \delta^2 \times 1 + \cdots = 5 + \frac{\delta}{1 - \delta}$$

如果采取诚信而不是欺骗，那么本期的收益为4(虽然小于5)，但在下一阶段还可得到完全相同的选择机会，如果始终选择诚信，那么其收益流的现值为

$$V = 4 + \delta 4 + \cdots = 4 + \delta(4 + \delta 4 + \cdots) = 4 + \delta V$$

$$V = \frac{4}{1 - \delta}$$

要保证参与者 j 选择诚信而不是欺骗，那么必须满足下列不等式

$$\frac{4}{1 - \delta} > 5 + \frac{\delta}{1 - \delta}$$

即贴现因子 $\delta \geq 1/4$ 。由此得到结论：当且仅当 $\delta \geq 1/4$ 时，在第一阶段，或者在前面结果都是(诚信，诚信)的 t 阶段，参与者 j 的最优反应为诚信。再加上前面已证明的，一旦某一阶段的结果偏离了(诚信，诚信)，参与者 j 的最优反应就是永远选择欺骗。因而给定参与者 i 的策略为触发策略，参与者 j 的最优策略同样为触发策略。利用对称性，可知给定参与者 j 的策略为触发策略，参与者 i 的最优策略同样为触发策略。这就证明了当且仅当贴现因子足够接近1($\delta \geq 1/4$)时，参与双方都采取触发策略是无限重复信用博弈的纳什均衡。

最后，证明触发策略是无限信用博弈的子博弈完美均衡。有时，要证明任意有限重复博弈或无限重复博弈中的纳什均衡为子博弈完美均衡是非常困难的，为了简化分析，我们需要一些辅助性结论。

在有限重复博弈 $G(T)$ 中，由第 $t+1$ 阶段开始的一个子博弈为 G 进行 $T-t$ 次博弈之后的重复博弈，可表示为 $G(T-t)$。重复博弈中从第一阶段到 t 阶段为止的每一种可能历史都对应着一个由第 $t+1$ 阶段开始的子博弈。与有限重复博弈类似，博弈 $G(\infty)$ 到阶段 t 为止有多少不同的可能历史，就有多少从 $t+1$ 阶段开始的子博弈 $G(\infty-t)$。

在重复两次的信用困境博弈中，第二阶段开始的子博弈有4个，而在重复两次的信用困境1和信用困境2中，第二阶段开始的于博弈则分别有9个和25个。在无穷重复信用困境博弈中，除了第一个阶段外，其余阶段开始的子博弈则可以分为两类：一是所有前而的历史都是(诚信，诚信)，二是所有前面的历史至

少出现了一次欺骗。

定义 5.4　单边背离特性(one-derivation property)在参与者 i 首先行动为起点的子博弈中,给定其他参与者的策略不变,参与者 i 的策略除了该子博弈的第一阶段外,其余阶段仍保持不变的情况下,在该子博弈的第一阶段参与者 i 改变行动(背离其策略)不可能增加其收益,就称该策略满足单边背离特性。

命题 5.2　有限重复博弈的一个策略组合为子博弈完美均衡,当且仅当其满足单边背离特性;在贴现因子 $\delta < 1$ 的情况下,无限重复博弈的一个策略组合为子博弈完美均衡,当且仅当其满足单边背离特性。

利用命题 5.2 可以极大地简化对触发策略为子博弈完美均衡的检验。在无限重复信用困境中,参与者 1 和参与者 2 是同时行动的,因而参与者 1 首先行动开始的子博弈就是任意阶段开始的子博弈。那么在无限重复信用困境博弈中,触发策略组合要满足单边背离特性,就要求在 $t>1$ 阶段开始的子博弈中,给定参与者 2 的策略为触发策略,参与者 1 的触发策略(除子博弈的第一阶段外)在其余阶段保持不变的情况下,参与者 1 在子博弈第一阶段改变行动不可能增加其收益。具体来讲:当 $\delta<1$ 时,当信用博弈进行到 t 阶段时,存在两个可能的历史过程,因而存在两类子博弈:① 所有以前阶段的结果都是(诚信,诚信)的子博弈;② 至少有一个前面阶段的结果不是(诚信,诚信)的子博弈。如果触发策略满足单边背离特性,则参与者 i 在第一类子博弈(即所有前面的结果都是(诚信,诚信))中改变行动(从诚信变为欺骗),并不能增加其收益。保持触发策略不变,贴现平均收益值为 4,而如果发生背离,那么在子博弈的第一阶段能得 5 的收益,但其后的所有阶段只能得到 1 的收益,这就要求 $4 \geq 5 + \delta/(1-\delta)$,即 $\delta \geq 1/4$。参与者在第二类子博弈中的最优策略是永远单纯重复阶段博弈的纳什均衡(欺骗,欺骗),参与者 i 改变行动(从欺骗变为诚信)并不能增加其收益。因为不背离的贴现平均收益值为 1,而背离为诚信在子博弈的第一阶段将得到 0,其后的所有阶段得 1,因而其贴现平均收益值为 δ,显然 $1>\delta$。综合上述两种情况,易知当 $1 > \delta \geq 1/4$ 时,触发策略是无限重复信用博弈任一子博弈的纳什均衡。这就证明了当 $\delta \geq 1/4$ 时,触发策略是无限重复信用困境的子博弈完美均衡。

㊥ 习题 5.3　*在无限重复信用困境中,不妨设参与者的策略为单边触发策略,即一开始选择诚信,如果对方始终诚信,那么选择诚信;只有前面的历史被人欺骗过,才会永远选择欺骗。请证明单边触发策略为无限重复信用困境的纳什均衡,但却不是子博弈完美均衡。*

(二) 有限惩罚策略

顾名思义,有限惩罚就是惩罚只会持续一段时间,而不会永远持续下去。有

限惩罚策略的正式定义如下：

一开始选择诚信，如 $t>1$ 期之前的历史都为诚信，那么在 t 期选择诚信；如果 $t-1$ 期出现欺骗，那么在随后的 k 期中都将选择欺骗，从 $t+k$ 期开始重新选择诚信。

有限惩罚策略可以用图 5－7 形象地表示出来，即如果发生欺骗，那么在随后的 1 到 k 期无论对手如何都将选择欺骗，然后又回到诚信。

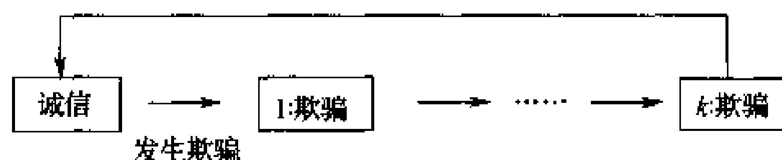


图 5－7　有限惩罚策略

我们不妨假设 $k=2$。要证明有限惩罚策略为无限重复信用博弈的纳什均衡需考虑如下几种情况：

第一，没有欺骗，参与者收益的现值为

$$V = 4 + 4\delta + 4\delta^2 + \cdots = \frac{4}{1-\delta}$$

第二，给定参与者 i 的有限惩罚策略，参与者 j 始终欺骗，其收益的现值为

$$V = 5 + \delta + \delta^2 + 5\delta^3 + \delta^4 + \delta^5 + 5\delta^6 + \cdots = \frac{5+\delta+\delta^2}{1-\delta^3}$$

如果诚信优于始终欺骗，那么就必须满足

$$4 \geqslant \frac{5+\delta+\delta^2}{1+\delta+\delta^2}$$

$$\delta \geqslant \frac{\sqrt{21}-3}{6}$$

第三，给定参与者 i 的有限惩罚策略，参与者 j 偶尔欺骗，不妨设发生了一次欺骗（不失一般性）。那么如果诚信优于偶尔欺骗，就要求

$$4 + 4\delta + 4\delta^2 \geqslant 5 + \delta + \delta^2$$

即 $\delta \geqslant \frac{\sqrt{21}-3}{6}$时，诚信要优于（永远或偶尔）欺骗，因而有限惩罚策略组合是重复信用博弈中的纳什均衡。要证明有限惩罚策略还是无限重复信用博弈任一子博弈的纳什均衡，将十分的困难和繁琐，但利用命题 5.2 就容易得多。

实际上，博弈进行到 $t-1$ 阶段可能的历史无非四种：① 没有欺骗，那么参与双方继续选择诚信是最优策略。② 刚好发生欺骗，那么参与双方连接选择两次欺骗，然后回到诚信是最优策略。③ 惩罚的第一期，那么最优选择就是接着

选择欺骗(因为对方会欺骗),然后选择诚信回到合作局面。④ 惩罚的第二期,那么接下来的最优选择就是回到诚信。所以当 $\delta \geqslant \frac{\sqrt{21}-3}{6}$ 时,在上面四类子博弈中有限惩罚策略仍是纳什均衡,这就最终证明了有限惩罚策略为无限重复信用困境的子博弈完美均衡。上面我们没有给出具体的数值证明,目的是想让读者自己来完成。

㊥ 习题 5.4 在无限重复信用困境中,参与者的策略为单边有限惩罚策略,即一开始选择诚信,如果对方没有欺骗过自己,那么选择诚信;如果被人欺骗,那么在随后的 k 期选择欺骗,然后重新选择诚信。请证明单边有限惩罚策略是纳什均衡,但却不是子博弈完美均衡。

(三) 一报还一报策略

一报还一报策略通俗地讲就是"以其人之道,还治其人之身。"其正式的定义如下:

在第一阶段选择诚信。在 $t>1$ 的任意阶段选择的策略等于对手在 $t-1$ 阶段选择的策略。

一报还一报策略也可以用图 5-8 形象地表示出来。(x,欺骗)表示对手选择欺骗,那么我也将选择欺骗;(x,诚信)则表示如果对手选择诚信,那么我将回到诚信的状态。x 代表我的行动是任意的。

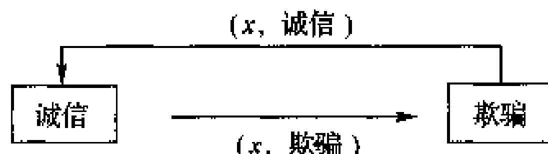


图 5-8 一报还一报策略

一报还一报策略最大的特点就是,如果参与者 j 欺骗了参与者 i,要使参与者 i 重新选择诚信,除非参与者 j 也被参与者 i 欺骗一次。

要就证明一报还一报策略是无限重复信用博弈的纳什均衡,需要考虑如下几种情况:

第一,没有欺骗,那么其贴现平均收益值为 4。

第二,存在欺骗,则有两种可能情况。一种是永远欺骗下去,那么其贴现平均收益值为 $5+4\delta$。另一种是开始欺骗,面后后悔,希望重新实现合作。由于我们考虑的是一报还一报策略,因而后悔的形式具有多种可能,这里我们不妨假设参与者 j 在第一阶段欺骗,但在第二阶段他开始后悔,从而选择诚信,以便在第三阶段重新实现合作,当然也可以继续选择欺骗,直到 t 阶段选择诚信,在 $t+1$

阶段重新获得合作。

第三，参与者 j 轮流选择欺骗和诚信。

那么如果在第二阶段后悔，则现值等于

$$V=5+\delta\times 0+4\delta^2+4\delta^3+\cdots=5+\frac{4\delta^2}{1-\delta} \qquad (5-T-1)$$

即参与者 j 在第一阶段欺骗得到 5 的收益，但在第二阶段选择诚信（参与者 i 选择欺骗，从而对自己进行惩罚，对参与者 i 进行一次补偿），得到收益 0，到了第三阶段重新实现合作。

另一种情况是参与者 j 在 $t>2$ 期才后悔，那么他在 t 期选择诚信，从而在 $t+1$ 阶段重新获得合作机会。则现值为

$$V=5+\delta+\delta^2+\cdots+0\times\delta^{t-1}+4\delta^t+\cdots$$

$$=5+\frac{\delta-\delta^{t-1}}{1-\delta}+\frac{4\delta^t}{1-\delta} \qquad (5-T-2)$$

最后一种情况就是参与者 j 轮流选择欺骗和诚信，那么其现值为

$$V=5+0+5\delta^2+0+5\delta^4+\cdots=\frac{5}{1-\delta^2}$$

由于(5－T－2)包含 δ^t，不方便与永远诚信的贴现收益平均值相比较，但容易证明当贴现因子 $\delta\geqslant 1/4$ 时，(5－T－1)式大于(5－T－2)式，即

$$4\delta^2+4\delta^3+\cdots+4\delta^{t-1}\geqslant\delta+\delta^2+\cdots+\delta^{t-2}$$

$$4\delta^2\frac{(1-\delta^{t-2})}{1-\delta}\geqslant\delta\frac{(1-\delta^{t-2})}{1-\delta}$$

(5－T－1)式大于等于(5－T－2)式说明后悔越早越好，因为 $t\geqslant 3$，所以上述结论对所有 t 期后悔的策略都成立，即它具有一般性。

因而如果一报还一报策略为纳什均衡，那么它必须同时满足下列各式：

$$4\geqslant 5-4\delta\rightarrow\delta\geqslant 1/4$$

$$4\geqslant 5(1-\delta)+4\delta^2\rightarrow 1\geqslant\delta\geqslant 1/4$$

$$4\geqslant\frac{5}{1+\delta}\rightarrow\delta\geqslant 1/4$$

即贴现因子 $\delta\geqslant 1/4$ 。因而可知，在给定参与者 i 的策略为一报还一报策略的情况下，一报还一报策略是参与者 j 的最优反应。利用对称性可知，给定参与者 j 的策略为一报还一报的情况下，参与者 i 的最优策略也是一报还一报。综上所述，当贴现因子 δ 接近 1（$\delta\geqslant 1/4$）时，一报还一报策略组合是重复无限信用博弈的纳什均衡。

当博弈进行到 t 阶段时，存在着四种可能的历史：

(1) $t-1$ 期的结果是（诚信，诚信），那么满足单边背离特性就要求 $\delta\geqslant 1/4$。

(2) $t-1$ 期的结果是(欺骗,欺骗),如果继续选择欺骗是最优选择,就需满足 $1\geqslant(1-\delta)(0+5\delta+0+5\delta^3+\cdots)\rightarrow 1/4\geqslant\delta$,即始终欺骗的贴现平均收益值要大于从欺骗背离到诚信,同时保持一报还一报策略不变时的贴现平均收益值。

(3) $t-1$ 期的结果是(诚信,欺骗),那么对于参与者 1 而言,从 t 期开始的子博弈的贴现值为 $5+0+5\delta^2+0+5\delta^4+\cdots$;如果背离选择诚信,那么贴现平均收益值为 4。这就意味着要防止背离,就必须满足 $1/4\geqslant\delta$。对于参与者 2 而言,从 t 期开始的子博弈的贴现值为 $0+5\delta+0+5\delta^3+\cdots$;如果背离选择欺骗,那么贴现平均收益值为 1。这就意味着要防止背离,必须满足 $\delta\geqslant 1/4$。

(4) $t-1$ 期的结果是(欺骗,诚信),分析同理,不再赘述。

综上所述,只有当 $\delta=1/4$ 时,一报还一报策略才是重复信用博弈的子博弈的纳什均衡,连同整个博弈纳什均衡 $\delta\geqslant 1/4$ 的要求,因而如果贴现因子 $\delta=1/4$,那么一报还一报策略组合就是无限重复信用困境的子博弈完美均衡。

㊂ 习题 5.5 将信用困境中的(诚信,诚信)组合下双方的收益改为 x,(诚信,欺骗)组合中,诚信方得 0,但欺骗方得 y,(欺骗,欺骗)组合下双方的收益为 z,它们之间的关系满足 $y-x=z, y\geqslant x\geqslant z>0$。那么请你证明,当 $\delta\geqslant z/x$ 时,触发策略为子博弈完美均衡。

实际上,根据定义,整个动态博弈本身就是一个子博弈,所以单边背离特性同样可以运用于整个博弈,而不仅仅是子博弈。这意味着,如果我们要证明某个策略组合为子博弈完美均衡,那么只需要检验它在整个博弈和所有可能的子博弈中是否满足单边背离特性,而无需证明其是整个重复博弈的纳什均衡。

例如,给定参与者采用一报还一报的策略组合,那么在无限重复信用困境的第一阶段,如参与者不背离,其贴现平均收益值等于 4,如背离,其贴现平均收益值就等于 $(1-\delta)(5+0+5\delta^2+0+5\delta^4+\cdots)=5/(1+\delta)$,所以当 $\delta\geqslant 1/4$ 时,一报还一报策略组合满足单边背离特性,与其为纳什均衡的条件完全一样。

上述分析都是在信用困境这一具体的博弈模型上进行的,一个重要的问题是无限重复信用博弈中得出的结论是否适用于一般性的无限重复博弈?幸运的是,答案是肯定的,这就引出了无名氏定理。

第三节 无名氏定理

在无限重复信用博弈中,实际上存在着各种各样的结果路径(全历史),每一个阶段都为(诚信,诚信)或每一个阶段都为(欺骗,欺骗)只不过是其中的一

个路径而已。例如,((诚信,诚信),(诚信,欺骗),(欺骗,欺骗),……)就是一个可能的结果路径。这个结果路径的贴现平均收益值对于参与者1而言,如果贴现因子δ充分接近1,一定位于1和4之间,因为对于参与者1,其贴现平均收益值为$4(1-\delta)+\delta^2$(参与者2的贴现平均收益值为$4+\delta-4\delta^2$)。显然一定存在一个$1>a>0$,使得$4a+1(1-a)=4(1-\delta)+\delta^2$。4是结果路径(诚信,诚信)的贴现平均收益值,而1则是结果路径(欺骗,欺骗)的贴现平均收益值。这说明对于任一结果路径的贴现平均收益值,可以用相应的参与者在G中的收益的凸组合来近似地表示。

定义5.5 设$\{x_1,\cdots,x_m\}$为策略式博弈G中所有可能的纯策略组合下相对应的收益组合的集合(注意其中的x为向量),$\{x_1,\cdots,x_m\}$的一个凸组合就称为G的一个**可行收益组合**,$\{x_1,\cdots,x_m\}$的所有凸组合集合就称为G的可行收益组合集合。所谓凸组合就是指$a_1x_1+\cdots+a_mx_m$,其中任意$a_i\geq 0$,并且$\sum_{i=1}^{m}a_i=1$。

例如,在信用困境博弈中,存在着4个纯策略组合,因而存在着4个相应的收益组合:(1,1),(0,5),(4,4)和(5,0),给定$a_i\geq 0$和$\sum_{i=1}^{4}a_i=1$,那么$a_1(1,1)+a_2(0,5)+a_3(4,4)+a_4(5,0)$就是一个凸组合。信用困境博弈中的可行收益组合集合即为图5-9中的阴影部分。可行收益组合集合包括$1<x<4$的所有(x,x),它们由(1,1)与(4,4)的凸组合得出($a_2=a_4=0$),以及满足$y+z=5$且$0<y<5$的所有(y,z),它们由(0,5)和(5,0)的凸组合得出($a_1=a_3=0$)。图5-9中阴影区域之内的其他组合则由两个以上的纯策略收益组合的凸组合得出。引入可行收益组合的意义在于:任一结果路径(对应着某个策略)的贴现平均收益值都等于某个可行收益组合。

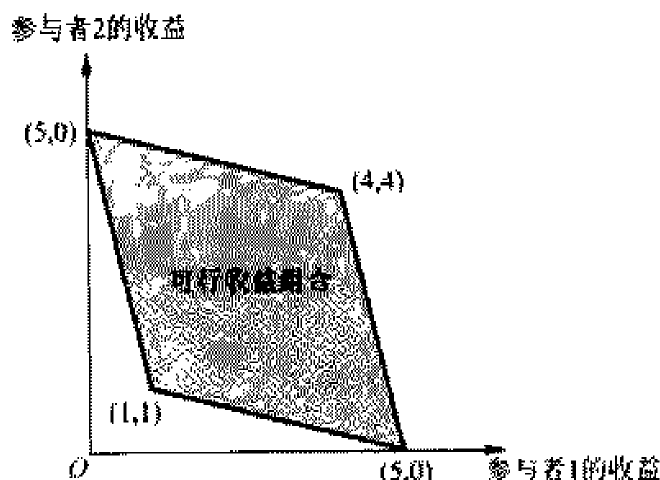


图5-9 信用困境的可行收益组合集合

理解可行收益的另一种方法就是参与者随机地选择策略(即混合策略,从而使得参与者无法判断下阶段对手可能采取的行动),参与者可以使用一个共同的随机发生器,例如,根据掷一枚硬币的结果选择(欺骗,诚信)或(诚信,欺骗),他们可得到的期望收益为就为(2.5,2.5)。

定理 5.1　无名氏定理(Friedman,1971)　$e=(e_1,\cdots,e_n)$为有限策略式博弈 G 的一个纳什均衡下的收益组合,$x=(x_1,\cdots,x_n)$则表示 G 的任一可行收益组合。如果 $x>e$,则存在一个 δ^*,使得对所有的 $\delta>\delta^*$,无限重复博弈 $G(\infty)$存在一个子博弈完美均衡,使得其结果路径的贴现平均收益值达到 x。

证:已知在无限重复博弈 $G(\infty)$中,设 $e=(e_1,\cdots,e_n)$为 G 的纳什均衡策略下的收益组合,并存在一个可行收益组合$(x_1,\cdots,x_n)$,并且 $x_i>e_i,i\in N$。我们要证明的就是,当贴现因子 $\delta\to1$ 时,存在一个合作性策略组合,使得$(x_1,\cdots,x_n)$成为子博弈完美均衡解的结果。我们把能够现实收益 x_i 的策略称为合作性策略。

我们不妨考虑触发策略:如果所有参与者都采用合作性策略,那么在每一阶段参与者 i 能得到的收益为 x_i(平均收益);如果有一个参与者在 t 期背叛了合作性策略,他在背叛期的收益为 $d_i>x_i$,那么 t 期以后的所有阶段,其他参与者将选择 G 中的纳什均衡。

我们要证明的就是触发策略是无限重复博弈 $G(\infty)$的子博弈完美均衡,并且$(x_1,\cdots,x_n)$作为完美解的结果出现。

在给定其他参与者的策略为触发策略的情况下,参与者 i 永远选择合作性策略的现值为

$$V=\frac{x_i}{1-\delta}$$

如果参与者 i 选择背叛性策略(非合作性的策略),那么根据触发策略,其他参与者将在随后的所有阶段选择 G 中的纳什均衡,那么现值为

$$V=d_i+\frac{e_i\delta}{1-\delta}$$

由于背叛前的收益一样,所以我们只要从第一期考虑即可。

要使合作解出现,必须要求

$$\frac{x_i}{1-\delta}>d_i+\frac{e_i\delta}{1-\delta}$$

整理后得

$$\delta>\frac{d_i-x_i}{d_i-e_i}$$

也就是说，当参与者 i 的贴现因子充分接近 1，那么参与者的最优选择就是选择合作性策略。当我们对其他参与者 $j\neq i$，重复以上过程，并且取

$$\delta > \max\left\{\frac{d_i - x_i}{d_i - e_i}\right\}_{i\in I}$$

那么触发策略就是无限重复博弈 $G(\infty)$ 的纳什均衡。

在触发策略下有两种可能的历史过程：第一种是在 t 期以前都是合作性策略，显然，t 期开始的子博弈中参与者拥有同原博弈一样的选择机会，子博弈等价于原博弈，可知触发策略是从 t 期开始的子博弈的纳什均衡；第二种是在 t 期之前出现过背叛性策略，根据触发策略，t 期及以后期的结果应当是两个参与者都会选择阶段博弈的纳什均衡（即非合作性策略），那么会有人在 t 期选择背离性的策略吗（即选择合作性策略）？显然不会，因为这样做收益为 0。这就证明了触发策略是无限重复博弈 $G(\infty)$ 中的一个子博弈完美均衡，并且能够实现贴现平均收益值 $(x_1,\cdots,x_n)$。命题得证。

在无限重复信用困境中，u（欺骗，欺骗）具有双重含义：① 对于参与者 i 而言，无论其采用什么策略，所得的收益都不可能小于 u（欺骗，欺骗），如果小于，他不如选择欺骗。② 对于参与者 j，则可以选择欺骗（比如为了惩罚参与者 i），从而使得参与者 i 的收益不可能超过 u（欺骗，欺骗）。那么我们如何将信用困境博弈中的 u（欺骗，欺骗）所具有的这一特征抽象出来呢？从参与者 i 的角度出发，给定参与者 j 的策略，其总是倾尽全力使其收益最大化，用数学来描述就是

$$\max_{s_i\in S_i} u_i(s_i,s_j)。$$

从参与者 j 的角度出发，则是倾尽全力使参与者 i 的收益最小化，即

$$\min_{s_j\in S_j}\{\max_{s_i\in S_i} u_i(s_i,s_j)\}。$$

定义 5.6　极小极大收益（minmax payoff）策略式博弈 $G=\{N,S,u\}$ 中任意参与者 $i(i\in N)$ 的极小极大收益为

$$\min_{s_{-i}\in S_{-i}}\{\max_{s_i\in S_i} u_i(s_i,s_{-i})\} \tag{5-4}$$

其中 $-i$ 代表除 i 以外的其他参与者。

极小极大收益实际上是一个博弈的参与者无论如何都能获得的收益，因而是一个保留收益[①]。对于存在合作可能的任一策略式博弈，纯纳什均衡收益绝对不可能小于极小极大收益，将纳什均衡的定义与(5-4)式比较就能证明这点。通常，一个博弈如果存在严格优策略，那么保留收益等于纳什均衡下

① 在零和博弈中，极小极大收益并不等同于保留收益，原因在于其是完全对抗性的。

的收益。在信用困境博弈中,1 既是一个极小极大收益,也是一个保留收益。在古诺模型中(见例 2.17),极小极大收益并不是纳什均衡下的利润,而是等于 0。

㊎ 习题 5.6 石头、剪刀、布博弈的极小极大收益是否是保留收益?如果不是的话,为什么?猎鹿博弈、伯川德模型的极小极大收益是多少?

一个有趣的问题是:如果我们采用更为严厉的惩罚措施,那么能否实现更高的合作收益?(纳什均衡不一定是最严厉的惩罚措施,而极小极大收益对应着的策略一定是更为严厉的惩罚措施)这一点被阿布勒(Abreu,1988)所证明,如果参与者采用更为严厉的可信的惩罚措施,那么能够实现比果用纳什均衡收益惩罚措施更高的贴现平均收益值。在信用困境博弈中,由于极小极大收益等于纳什均衡收益,因而(4,4)是无限重复信用博弈所能实现的最高贴现平均收益值。如果将定理 5.1 中的纳什均衡收益替换成极小极大收益,那么无名氏定理是否仍然成立?

定理 5.2 无限重复 n 人无名氏定理 $y=(y_1,\cdots,y_n)$ 为有限策略式博弈 G 的极小极大收益组合,$x=(x_1,\cdots,x_n)$ 则表示 G 的任一可行收益组合。如果 $x>y$,则存在一个 δ^*,使得对所有的 $\delta>\delta^*$,无限重复博弈 $G(\infty)$ 存在一个纳什均衡,使得其结果路径的贴现平均收益值达到 x。

定理 5.2 与定理 5.1 的根本区别就在于子博弈完美均衡变成了纳什均衡,原因在于对于某些可行收益组合 $x(e>x>y)$,在 n 人博弈($n>2$)中如果有两人的偏好相同,那么无论 δ 如何取值,有可能不存在子博弈完美均衡,而只存在纳什均衡。如果 n 人博弈中,没有哪个参与者的偏好与其他参与者的偏好一样,或没有哪个参与者的收益函数等价于其他参与者的收益函数①,那么定理 5.2 中的纳什均衡就可以替换为子博弈完美均衡(Abreu,Dutta and Smith,1994)。这里的等价是指线性变换,即在阶段博弈 G 中,$u_i(a)=c+bu_j(a)$,其中 c 为任一实数,$b>0$。如果 $c=0,b=1$,那么两人的收益函数完全相等。此外,富登伯格(Fudenberg)和马斯金(Maskin)1986 年证明,对于两人博弈,将纳什均衡收益替换为极小极大收益,定理 5.1 仍然成立。

定理 5.3 无限重复 2 人无名氏定理 $y=(y_1,y_2)$ 为两人有限策略式博弈 G 的极小极大收益组合,$x=(x_1,x_2)$ 则表示 G 的任一可行收益组合。如果 $x>y$,则存在一个 δ^*,使得对所有的 $\delta>\delta^*$,无限重复博弈 $G(\infty)$ 存在一个子博弈完美均衡,使得其结采路径的贴现平均收益值达到 x。

---

① 如果引人有效极小极大收益的概念,那么就可以放弃这个条件,使得无名氏定理对子博弈完美均衡成立,见 Wen(1994)的论文。

为了说明确实存在一个子博弈完美均衡，使得结果路径的贴现平均收益值小于纳什均衡收益，但大于极小极大收值，不妨考虑图 5－10 所示的无限重复博弈。

		参与者2 A	参与者2 B	参与者2 C
参与者1	A	4,4	3,0	1,0
参与者1	B	0,3	2,2	1,0
参与者1	C	0,1	0,1	0,0

图 5－10　极小极大收益小于纳什均衡收益

图 5－10 所示博弈的纳什均衡收益为(4,4)，极小极大收益为 1。现在我们所想证明的就是存在一个策略组合(s,s)，使得该博弈每一期的结果都是(B,B)，显然(2,2)是该博弈的一个可行收益组合。我们不妨假设策略 s 为有限惩罚策略，即开始选择 B，如果上一期是(B,B)，那么选择 B；如果上一期的结果不是(B,B)，那么在随后的两期选择 C，只有连续两期的结果都为(C,C)，才重新选择 B，否则仍选择 C。

㊐ 习题 5.7　请证明上面的有限惩罚策略是图 5－10 所示无限重复博弈的纳什均衡。

现在我们证明有限惩罚策略在子博弈中满足单边背离特性。显然存在三种可能的历史：

(1) 上一期的状态为(B,B)，如果坚持有限惩罚策略，那么后三期的收益为(2,2,2)；如果发生背离，那么收益为(3,0,0)，因而只要 $\delta \geqslant (\sqrt{3}-1)/2$，就满足单边背离特性。实际上，这时参与者拥有同原博弈一样的选择权，子博弈等于原博弈 $G(\infty)$，所以有限惩罚策略在该子博弈仍为纳什均衡。

(2) 上一期刚好发生非(B,B)的状态，子博弈的开始阶段为惩罚的第一期，如果坚持有限惩罚策略，那么其在后三期的收益为(0,0,2)，而如果背离，那么为(1,0,0)，因而只要 $\delta \geqslant \sqrt{2}/2$，就满足单边背离特性。

(3) 子博弈的开始阶段为惩罚的第二期，如果坚持有限惩罚策略，那么其后三期的收益为(0,2,2)，如果背离，那么为(1,0,0)，因而只要 $\delta \geqslant (\sqrt{3}-1)/2$，就满足单边背离特性。

综上所述，当 $\delta \geqslant \sqrt{2}/2$ 时，有限惩罚策略为图 5－10 所示无限重复博弈的子博弈完美均衡，其形成的结果路径为(B,B)，相应的收益为(2,2) < (4,4)的纳什均衡收益。

前面的分析都是基于无限重复博弈，那么在博弈是有限重复的情况下，是否也存在着类似的无名氏定理呢？在第二节中我们曾提到多重均衡，如果情况更为一般，那么是否也能在有限重复博弈中实现合作解？比诺维特和克瑞史纳（Benoit and Krishna，1985，1987）证明如果存在纳什均衡收益大于极小极大收益，那么如果 T 足够长，就存在一个纳什均衡使得每一个参与者的平均收益无限接近 $x(>y)$，其中 x 为可行收益组合，y 为极小极大收益组合。

定理 5.4　有限重复 n 人无名氏定理 G 为有限策略式博弈，如果① G 存在一个纳什均衡使得每一个参与者的纳什均衡收益都大于他的极小极大收益，即 $e>y$；② 令 x 为大于极小极大收益组合 y 的一个可行收益组合，即 $x>y$；那么就存在一个 T^*，以至于对于所有的 $T>T^*$ 和 $\delta=1$，$G(T)$ 存在一个纳什均衡，使得该结果路径下每一个参与者的平均收益值（因为 $\delta=1$，所以不是贴现平均收益值）无限接近 x，用公式表示就是

$$\left|\frac{1}{T}\sum_{t=1}^{T}u_i(a^t)-x_i\right|<\varepsilon$$

对所有的 $i\in N$ 成立。其中 a^t 表示 $G(T)$ 中 t 阶段的行动组合，$\varepsilon>0$。

定理 5.5　有限重复 2 人无名氏定理 G 为 2 人有限策略式博弈，如果① G 存在两个纳什均衡，并且对每一个参与者而言，其两个纳什均衡收益都不相等；② 令 x 为大于极小极大收益组合 y 的一个可行收益组合，即 $x>y$；那么就存在一个 T^*，以至于对于所有的 $T>T^*$ 和 $\delta=1$，$G(T)$ 存在一个子博弈完美均衡，使得该结果路径下每一个参与者的平均收益值（因为 $\delta=1$，所以不是贴现平均收益值）无限接近 x。

㊁ 习题 5.8　我们已经知道重复两次的信用困境 1 和信用困境 2 存在子博弈完美均衡，其实现的可行收益大于极小极大收益。现在请你证明重复 $k>2$ 次的信用困境 1 和信用困境 2 是否满足定理 5.4 和定理 5.5。

第四节　经典举例

为了巩固和加深理解，现在我们就举一些经典的例子。

一、无限重复古诺双头垄断下的共谋

在第二章中曾分析过静态下的古诺双头垄断模型，得出的结论是古诺模型中垄断企业间的共谋（即合作）是不稳定的，共谋下的产量 $(a-c)/4$ 不可能作为均衡结果而出现。但在无限重复博弈下，根据无名氏定理可知，当 δ 接近 1 时，

一定存在一个子博弈完美均衡使得共谋下的产量在无限重复古诺模型中出现。这就解释了为什么现实中会出现反托拉斯法。即便存在反托拉斯法，垄断行业仍然存在普遍的非公开勾结。

首先将第二章第四节中的古诺双头垄断模型的主要内容复述如下：市场中的总产量为两个厂商产量之和 $Q = q_1 + q_2$，市场的总需求曲线为 $P(Q) = a - Q$，假定 $a > Q$。每一个企业的边际成本为 c，固定成本为 0，两企业同时选择产量。唯一的纳什均衡为每一企业选择 $(a-c)/3$ 的产量，不妨称之为纳什均衡产量并用 q_c 表示。由于均衡条件下的总产量 $2(a-c)/3$ 大于卡特尔垄断产量 $q_m = (a-c)/2$，如果两企业平均分配卡特尔产量，即 $q_i = q_m/2$，那么每一企业能获得较纳什均衡产量更大的利润。

现在考虑以上述古诺模型为阶段博弈的无限重复博弈，两企业的贴现因子均为 δ。不妨假设两个垄断企业采取的是触发策略：

在第一阶段生产卡特尔产量的一半，$q_m/2$。第 t 阶段，如果前面 $t-1$ 个阶段两个企业的产量都为 $q_m/2$，则生产 $q_m/2$；否则，生产纳什均衡产量 q_c。

由于证明过程与前面证明信用博弈的思路完全一样，这里只把主要的思想简单叙述一遍。

当双方都生产 $q_m/2$ 时，每个企业的利润为 $(a-c)^2/8$，用 π_m 来表示。当双方都生产 q_c 时，每个企业的利润为 $(a-c)^2/9$，用 π_c 表示。最后，如果企业 i 将在本期生产 $q_m/2$，而企业 j 背叛卡特尔产量，那么本期利润最大化的产量为下式的解

$$\max_{q_j}\left(a - q_j - \frac{q_m}{2} - c\right)q_j$$

它的解为 $q_j = 3(a-c)/8$，相应的利润水平为 $9(a-c)^2/64$，用 π_d 来表示。那么，要使上述触发策略成为无限重复古诺模型的子博弈完美解，就必须满足

$$\frac{\pi_m}{1-\delta} > \pi_d + \frac{\delta\pi_c}{1-\delta} \qquad (5-C-1)$$

将 π_m, π_c, π_d 的值分别带入(5－C－1)式可得 $\delta > 9/17$。由此可知，当 $\delta > 9/17$ 时，上述触发策略将是无限重复古诺模型的子博弈完美均衡。

一个有趣的问题是，如果 $\delta \leq 9/17$，那么触发策略还能不能维持一个“理想”的产量。由于纳什均衡利润 $\pi_c = (a-c)^2/9$，卡特尔利润 $\pi_m = (a-c)^2/8$，根据无名氏定理，在 $[(a-c)^2/9, (a-c)^2/8]$ 之间存在着许多的可行利润，因面一定存在某些子博弈完美均衡能够实现这些可行利润。见图 5－11。

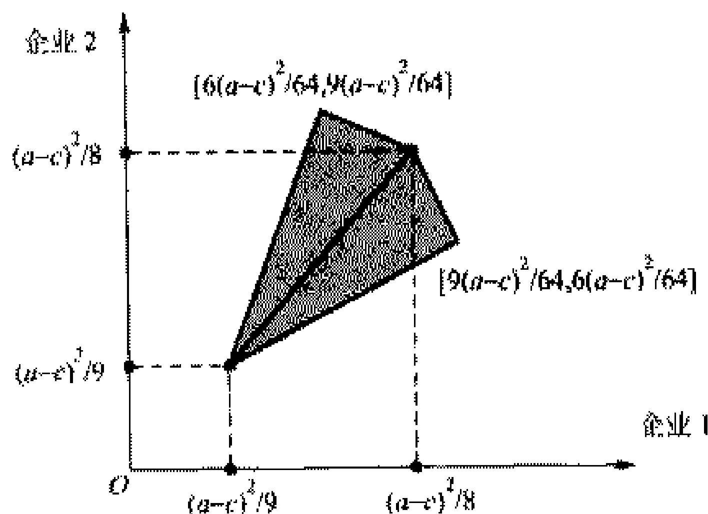


图 5-11 可行利润集合

有两种思路来分析这一问题。首先计算对任意一个给定的 δ 值，如果双方都采用触发策略，即一旦出现背离就永远转到纳什均衡产量，企业可以达到的利润最大化产量。因为当企业的产量小于 $q_m/2$ 时，是一个严格劣策略（想一想在对手的产量为 0 时，每个企业利润最大化下的产量是多少？给定对手生产 $(a-c)/2$，那么企业的最优产量选择又是多少？），所以企业的产量选择不可能小于 $q_m/2$；其次，对于任意的 δ 值，永远简单重复纳什均衡产量显然是一个子博弈完美均衡。因而，每一个企业所能选择的产量一定位于区间 $[q_m/2, q_c]$ 之内。为了计算这一产量，考虑如下的触发策略：

第一阶段生产 q^*。在第 t 阶段，如果在此之前的 $t-1$ 个阶段两企业的产量都是 q^*，生产 q^*；否则，生产纳什均衡产量 q_c。

如果双方都生产 q^*，每个企业的利润为 $(a-2q^*-c)q^*$，用 π^* 来表示。如果企业 i 计划在当期生产 q^*，而企业 j 背叛，那么当期利润最大化的产量为下式的解

$$\max_{q_j}(a-q^*-q_j-c)q_j$$

其解为 $q_j=(a-q^*-c)/2$，相应的利润为 $(a-q^*-c)^2/4$，仍用 π_d 表示。当下式成立时，两个企业都采取上面给出的触发策略为完美均衡

$$\frac{\pi^*}{1-\delta} > \pi_d + \frac{\delta\pi_c}{1-\delta}$$

即

$$q^* = \frac{9-5\delta}{3(9-\delta)}(a-c)$$

它随 δ 单调递减，且当 $\delta=9/17$ 时，达到 $q_m/2$，当 $\delta=0$ 时达到 q_c。

第二种思路就是企业威胁使用最严厉的可置信惩罚。企业的惩罚越严厉，那么对 δ 充分接近 1 的要求就越松。这里我们只简单地证明，如果惩罚更为严厉，那么在 $\delta=1/2$ 时，也可以达到卡特尔产量。考虑下面的萝卜加大棒(carrot and stick)策略：

在第一阶段生产卡特尔产量的一半，$q_m/2$。第 t 阶段，如果两个企业在第 $t-1$ 阶段都生产 $q_m/2$，则生产 $q_m/2$；如果两个企业在 $t-1$ 阶段的产量都是 x，则生产 $q_m/2$；其他情况下生产 x。

这一策略为参与者提供了两种手段：其一是单阶段的惩罚，这时企业生产 x；其二是潜在无限阶段的合作，这时企业的产量为 $q_m/2$。如果任何一个企业偏离了合作，则惩罚开始。如果任何一个企业背离了惩罚，则会使博弈进入又一轮惩罚。如果两个企业都不背离惩罚，则在下一阶段又回到合作。

如果两企业都生产 x，那么每个企业的利润为 $(a-2x-c)x$，用 $\pi(x)$ 表示。令 $V(x)$ 表示当期的利润为 $\pi(x)$，以后每阶段的利润永远是卡特尔利润的一半，企业总收益的现值

$$V(x) = \pi(x) + \frac{\delta\pi_m}{1-\delta}$$

如果企业 i 选择产量 x，而企业 j 背叛，那么企业 j 利润最大化时的产量为下式的解

$$\max_{q_j}(a-q_j-x-c)q_j$$

其解为 $q_j=(a-x-c)/2$，相应的利润为 $(a-x-c)^2/4$，用 $\pi_{dp}(x)$ 来表示。

为了便于大家理解，我们将所有涉及的产量和利润列成表 5-1。

表 5-1　产量和利润

产量	利润
$q_c=(a-c)/3$	$\pi_c=(a-c)^2/9$
$q_m=(a-c)/4$	$\pi_m=(a-c)^2/8$
$q_d=3(a-c)/8$	$\pi_d=9(a-c)^2/64$
x	$\pi(x)=(a-2x-c)x$
$q_{dp}=(a-x-c)/2$	$\pi_{dp}(x)=(a-x-c)^2/4$

如果两个企业都采用上面的萝卜加大棒策略，则无限重复古诺模型中进行到第 t 阶段的历史过程(子博弈)就可归为两类：① 合作的子博弈，其前面一个

阶段的结果是$(q_m/2,q_m/2)$或(x,x);② 惩罚的子博弈,其前面一个阶段的结果既非$(q_m/2,q_m/2)$,又不是(x,x)。两企业都采取上面的萝卜加大棒策略要成为一个子博弈完美均衡,则该策略在每一类子博弈中必须是纳什均衡。具体地说,第一,在合作子博弈中,选择永远合作要优于背叛之后的再次合作,即每一企业获得卡特尔利润的现值要大于本期得到π_d与下期得到惩罚的现值收益$V(x)$之和。

$$\frac{\pi_m}{1-\delta} > \pi_d + \delta V(x) \qquad (5-C-2)$$

将$V(x)$代入(5-C-2)式可得

$$(1+\delta)\pi_m > \pi_d + \delta\pi(x)$$

$$\delta[\pi_m - \pi(x)] > \pi_d - \pi_m \qquad (5-C-3)$$

它表示,在本期背离所得到的好处必须小于下一期惩罚带来损失的现值。实际上,上述结论非常直观,因为无论是永远合作还是背叛之后的再次合作,从第二期开始它们的现值都是一样的,所以是否背叛就取决于本期背叛所得好处是否大于下一期惩罚带来的损失现值。

将表5-1中的值相应带入(5-C-3)式,整理后得

$$8(1+\delta) > 9 + \frac{64\delta x}{a-c} - \frac{128\delta x^2}{(a-c)^2}$$

令$y=x/(a-c)$,$\delta=1/2$,得

$$64y^2 - 32y + 3 \geq 0$$

解得的结果是:如果$x/(a-c)$不在1/8到3/8之间,(5-C-2)式即可满足。

第二,在惩罚子博弈中,企业本期接受惩罚比起继续背叛,且下一期继续接受惩罚要好:

$$V(x) > \pi_{dp}(x) + \delta V(x) \qquad (5-C-4)$$

将$V(x)$代入(5-C-4)式,得

$$\delta[\pi_m - \pi(x)] > \pi_{dp}(x) - \pi(x) \qquad (5-C-5)$$

即老老实实接受惩罚要好过继续背叛。将表5-1中的值代入(5-C-5),仍令$\delta=1/2$,解得的结果是:如果$x/(a-c)$处于3/10到1/2之间,那么(5-C-4)式必然成立。从而,对贴现因子$\delta=1/2$可达到卡特尔产量的萝卜加大棒策略成为子博弈完美均衡的条件是$3/8<x/(a-c)<1/2$。

实际上,我们这里考察的古诺模型相对简单,它还可以进一步扩展。例如,引入状态的随机性,这时垄断企业面对的市场需求随着经济扩张和收缩随机波动;或是引入一定的信息不完全,其中一个重要的模型就是不完美监督模型。下

一个例子效率工资就会涉及这一点。

从某种意义上说,萝卜加大棒政策要比触发策略为优,当人们无法完美地观察到企业的背离时,情况就更是如此。例如企业不能观察到彼此的产量选择,只能观察到市场出清价格,而这一价格在每一阶段又会受到一些随机因素的冲击,例如经济衰退、战争。在这样的条件下,企业就很难判断市场出清价格下降是由于其他企业背叛共谋产量,还是由于其他不利因素所造成。根据触发策略,任何低于触发水平的价格都会引发价格战,并导致永远的惩罚。然而,市场因素一次严重的不利冲击也会使价格降到触发水平,从而导致永远的惩罚,由于惩罚是由于偶然因素引起,因而触发策略中导致惩罚永远进行下去的做法就不再是最优的,显然萝卜加大棒的政策要更为合理。

二、效率工资

在新凯恩斯宏观经济学中,效率工资是说明工资刚性和失业并存的重要模型。在效率工资模型中,企业为了激励工人努力工作,一方面支付比劳动力市场均衡为高的工资率,另一方面又威胁如果发现工人偷懒将立即开除。这种做法导致的后果就是工资刚性和人为制造的非自愿失业并存,并且这种状态能够在劳动力市场中长期存在。见图 5-12。

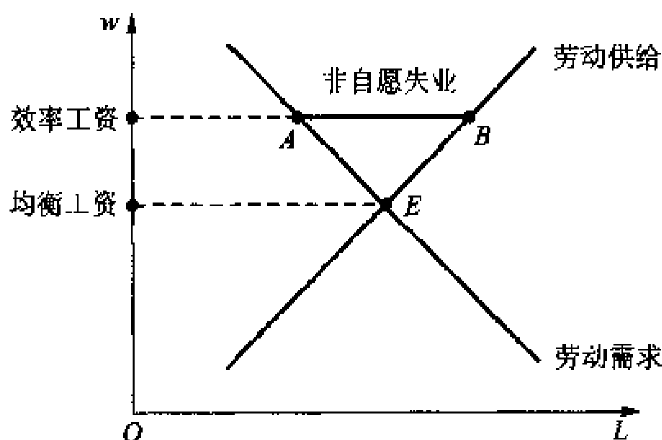


图 5-12　效率工资与失业

显然这种情况在传统经济学中是不可能出现的。然而我们在后面将证明,效率工资这种策略(实际上是一种触发策略)在企业和工人组成的无限重复博弈中却可以是一个子博弈完美均衡解。效率工资越高,失业工人的人数越多,一个被开除的工人寻找到新的就业机会的可能性就越小,于是开除的威胁就越有效。在均衡条件下,工资水平 w 和失业率 u 恰好可以使工人不敢偷懒,并且企业在工资水平 w 时的劳动需求恰好使得失业率等于 u。这里我们分析一个最简

单的模型[①],即只包含一个企业和一个工人。

考虑如下的阶段博弈:第一,企业给出一个工资水平 w;第二,工人根据 w 决定接受或拒绝工作。如果工人拒绝,则工人成为自我雇佣者,工资水平为 w_0,如果工人接受了 w,则工人选择是努力工作(会带来 e 的负效用)还是偷懒(不会带来任何负效用)。工人的努力程度企业无法观测,但企业能观测到工人的产出水平。产出可能高也可能低,为简单起见,不妨假设低产出为0,高产出为 $y>0$。如果工人努力工作则肯定可以得到高产出,但如果工人偷懒则以 p 的概率得到高产出,$1-p$ 的概率得到低产出。从而,在此模型中,低产出是偷懒无可辩驳的证据。

假设企业以 w 的工资雇佣了工人,那么如果工人努力工作,带来高产出时参与人的收益分别为:企业 $y-w$,工人 $w-e$。如果工人偷懒,则 e 变为0;如果出现低产出,则 y 变为0。假定 $y-e>w_0>py$,从而对工人来讲,受雇于企业并且努力工作是最优的,工人自我雇佣要优于受雇于企业并偷懒。

这一阶段博弈的纳什均衡解是企业开出0的工资,工人选择自我雇佣。工人受雇于企业这样的合作解并不会出现。因为如果企业开出 w 的工资,工人选择偷懒是一个严格优策略(选择努力收益为 $w-e<w-0$ 选择偷懒)。在有限重复博弈中结论同样成立。但如果这个博弈重复进行无穷多次,那么根据无名氏定理,一定存在一个子博弈完美均衡使得合作解,即工人受雇于企业并努力工作在重复博弈中出现,而效率工资就是该博弈的一个子博弈完美均衡。

如果所有前面的工资开价都是 w,所有的开价都被接受了,并且所有前期的产出都是高的,就称博弈的过程是"高工资、高产出"。企业的策略为第一阶段开出工资水平 $w=w^*$,并且在其后的每一阶段,如果博弈的过程是高工资、高产出,则继续开出工资水平 w;但其他情况下开出 $w=0$。工人的策略为如果 $w\geqslant w_0$,则接受企业的工资(否则,选择自我雇佣),并且如果博弈的过程(包括本阶段的工资)是高工资、高产出,则努力工作(否则偷懒)。

企业的策略类似于前面分析的触发策略:如果在前面所有阶段的博弈中都相互合作就继续合作,但一旦有一次合作被打破,就永远转向阶段博弈的纳什均衡解。工人的策略也类似于触发策略,但由于工人在阶段博弈中后行动,其策略也更加灵活一些。在一个基于静态阶段博弈的重复博弈中,一方的背离只能在当前阶段结束时才可观测到;但在基于动态阶段博弈的重复博弈中,首先行动方的背离在同一阶段就可被观测到,并可对其作出反应。工人的策

① Shapiro C. and Stiglitz J. "Equilibrium Unemployment as a discipline Device". *American Economic Review*, 1984, 74: 433–44

略将会是:如果前面所有阶段的博弈都合作则继续合作,如果企业一旦有背离就选择自己本阶段的最优行动,因为他知道将来所有阶段的博弈都将出现阶段博弈的纳什均衡。具体地说,如果 $w \neq w^*$,但 $w \geq w_0$,则工人将接受企业的工资但选择偷懒。

下面将导出上述双方的策略为子博弈完美均衡的条件。同前面一样,论证由两部分组成:① 导出双方策略成为纳什均衡的条件,② 证明它们是子博弈完美的。

假设企业在第一阶段开出的工资为 w^*。给定企业的策略,工人接受这一工资水平是最优的。如果工人努力工作,则他肯定得到高产出,那么企业将再次开出工资水平 w^*,而工人将在下一阶段就努力与否进行相同的决策。从而,如果对工人来讲努力工作是最优的,则工人收益的现值为

$$V_e = w^* - e + \delta V_e$$

如果工人偷懒,则工人将以 p 的概率得到高产出,这时下一阶段他还可以就努力与否进行决策;但工人还将以 $1-p$ 的概率得到低产出,这时企业将在以后永远开出工资 $w=0$,于是工人亦将永远选择自我雇佣。从而,如果对工人来讲偷懒是最优的,则工人的收益现值为

$$V_s = w^* + \delta\left[pV_s + (1-p)\frac{w_0}{1-\delta}\right]$$

$$V_s = \frac{1}{(1-p\delta)(1-\delta)}[(1-\delta)w^* + \delta(1-p)w_0]$$

对于工人来讲,要使选择努力为最优必须满足 $V_e \geq V_s$,即

$$w^* \geq w_0 + \frac{1-p\delta}{\delta(1-p)}e = w_0 + \left[1 + \frac{1-\delta}{\delta(1-p)}\right]e \quad (5-C-6)$$

于是,为激励工人努力工作,企业支付给工人的工资,不仅应补偿工人自我雇佣时的机会收入以及努力工作带来的负效用 w_0+e,而且还应包含工资升水 $(1-\delta)e/\delta(1-p)$。很自然,如果 p 接近于1(如果偷懒很难发现)则工资升水必须非常高才可以激励工人努力工作。另一方面,如果 $p=0$,则在下式成立时,工人努力工作是最优的

$$\frac{1}{1-\delta}(w^* - e) \geq w^* + \frac{\delta}{1-\delta}w_0$$

上式又可化为下式

$$w^* \geq w_0 + \left(1 + \frac{1-\delta}{\delta}\right)e$$

它实际上就是(5-C-6)式中 $p=0$ 时的情景。

(5-C-6)式成立只表明工人的策略为企业策略的最优反应,还需研究企

业支付 w 是否值得。给定工人的策略，企业在第一阶段的问题可归为进行以下选择：① 支付 $w=w^*$，并通过威胁工人一旦出现低产出就将其开除来激励工人努力工作，这样每一阶段都可得到 $y-w$ 的收益；② 支付 $w=0$，促使工人选择自我雇佣，自己在每一阶段的收益均为0。于是，企业策略成为工人策略最优反应的条件为

$$y-w^* \geqslant 0 \tag{5-C-7}$$

前面我们假设工人选择受雇于企业并努力工作是有效率的，即 $y-e>w_0$。但要使双方策略成为子博弈完美均衡，还需要进一步的条件：将(5-C-6)式和(5-C-7)式合并得

$$y-e \geqslant w_0+\frac{1-\delta}{\delta(1-p)}e$$

这说明要使效率工资能够成为完美均衡，即维持企业和工人的雇佣关系，需要一个充分接近1的贴现因子 δ，只有 δ 足够大，这样的合作关系才可能出现。

到此为止，我们已证明如果(5-C-6)和(5-C-6)成立，则前面给出的策略为纳什均衡。为证明这些策略同时又是子博弈完美的，首先需定义原重复博弈的子博弈。从前面的分析可知，如果阶段博弈为静态博弈，那么重复博弈的子博弈由某个阶段博弈开始，但在阶段博弈为序贯行动的情况下，子博弈不仅可从阶段博弈开始，还可以始于每一阶段之中——在工人观测到企业给出的工资水平之后。给定参与者的策略，我们可以把子博弈归为两类：① 始于高工资、高产出之后的子博弈；② 非高工资、高产出历史过程之后的子博弈。我们已证明高工资、高产出的历史过程下，参与者的上述策略为纳什均衡，剩下的就是要证明非高工资、高产出历史过程下的情况：由于工人将来不再会努力工作，企业促使工人选择自我就业是最优的；由于企业将在下一阶段及其后永远支付工资 $w=0$，工人在当前阶段的最优行动是偷懒，并只有在 $w>w_0$ 时才会接受给付的工资。因而，触发策略是无限重复效率工资模型中的子博弈完美均衡。

在效率工资模型中，一旦发生背离，自我雇佣就是永久性的。如果工人曾有一次被捉住偷懒，则企业在其后将永远给付工资 $w=0$；如果企业曾偏离 $w=w^*$（最优工资），则工人将不再努力工作，于是企业也不会再雇佣这个工人。同永远选择阶段博弈的纳什均衡解相比，双方参与者可能更愿意回到无限重复博弈的高工资、高产出均衡去，因而双方都有重新进行谈判的动机。但问题是，如果参与者知道惩罚将不会被执行，那么这种威胁就是不可置信的，以这种惩罚相威胁促成的合作便不再是一个均衡。

在存在劳动力市场的情况下，如果企业同时雇佣了许多工人，则它更倾向于不进行重新谈判，因为和一个工人的重新谈判会使处于高工资、高产出均衡的其

他工人(或将要选择这一均衡的工人)十分失望而改变策略。如果存在许多企业,问题是企业 j 是否会雇佣以前企业 i 曾解雇过的工人,合理的结果可能是"否"。因为它担心会使现有高工资、高产出的工人失望,正如一个企业的情况一样。诸如此类的原因可以解释为什么日本大企业间白领男性的成年雇员缺乏流动性。

换一种情况,如果被解雇的工人总可以找到比自我雇佣更好的工作,则这时那些新工作的工资(减去努力带来的负效用)就起到了自我雇佣收人 w_0 的作用(机会收人)。在一个被雇佣工人根本不会受到任何损失的极端情况下(例如他有许多工作可做,或是吃大锅饭),即使是无限重复博弈也无法提供对偷懒有效的惩罚,从而不存在工人将努力工作的子博弈完美均衡。效率工资模型实际上具有一般性,例如资本市场的债权人—债务人,股市的股东—公司经理人,以及公民—政府官员,如果债务人、经理人、政府官员具有许多的机会,它们的命运不是由委托人说了算(债务人复杂的社会关系、经理人的行政任命、政府官员的上级任命),那么代理人努力工作就不会是完美均衡解。

三、无通货膨胀的货币政策

考虑如下动态博弈,其中雇主和工人就名义工资进行谈判,之后货币当局选择货币供给,货币供给量又决定了通货膨胀率。雇主和工人在决定工资时都将尽力去预测通货膨胀。不过,一旦名义工资确定,真实的通货膨胀率如高于预期的通货膨胀率,将会使实际工资下降,导致雇主扩大雇佣人数,产出增加。这样货币当局就需要在通货膨胀成本和高通货膨胀使失业率降低、总产出提高之间进行权衡。

我们用下面的动态博弈来分析这一问题的简化形式[①]。首先,雇主和工人都希望能正确地预期到通胀,为了简化,不妨假设雇主和工人拥有的信息完全一样,因而雇主和工人的通胀预期值都是 π^e。这样就只需要考虑雇主(或是工人)即可,不妨设雇主的收益为 $-(\pi-\pi^e)^2$。第二,货币当局观测到这一预期并决定实际的通胀率 π。货币当局既希望通胀率越小越好,又希望实际产出 y 达到充分就业水平 y^*。可以把货币当局的收益用下式表示

$$U(\pi,y) = -c\pi^2 - (y-y^*)^2$$

其中的参数 $c>0$,c 越大说明通胀率目标越重要。假设实际产出可表示为如下的目标产出和实际通胀率的函数

① Barro R. and Gordon D. "Rules, Discretion, and Reputation in a Model of Monetary Policy". *Journal of Monetary Economics*. 1983, 12:101 – 121

$$y = by^* + d(\pi - \pi^e)$$

其中，$1 > b > 0$ 表明存在自然失业（从而如果没有意料外的通胀，则实际产出小于充分就业的产出水平），$d > 0$ 表示实际通胀率大于预期通胀率时，实际工资下降，因而实际产出增加。由此可以将货币当局的收益重新表示为

$$W(\pi, \pi^e) = -c\pi^2 - [(b-1)y^* + d(\pi - \pi^e)]^2$$

给定雇主预期的通胀率 π^e，货币当局选择 π 令上式最大化，即

$$\max_{\pi} W(\pi, \pi^e)$$

解一阶条件，得

$$\pi^*(\pi^e) = \frac{d}{c + d^2}[(1-b)y^* + d\pi^e] \qquad (5-M-1)$$

对于雇主而言，最优的结果就是正确预测 π^*，因而雇主的最优选择就是令 $\pi^e = \pi^*(\pi^e)$，代入(5 - M - 1)可解得

$$\pi^* = \frac{d(1-b)}{c}y^* = \pi^e$$

即（$\pi^e = \pi^*$，$\pi^* = d(1-b)y^*/c$）为该动态博弈的子博弈完美均衡。在这一子博弈完美均衡解中，实际通胀率大于0，但事实上，如果雇主持有理性预期（即 $\pi^e = \pi$），零通胀率不仅使货币当局的收益达到最大，而且也将使社会的福利达到最大（货币当局、雇主、工人三方的收益之和）。但这一结果，在有限重复货币政策模型中根本无法实现。

现在考虑雇主和货币当局的贴现因子都为 δ 时的无限重复博弈的情况。我们将证明：如果双方都采用触发策略，那么它将成为子博弈完美均衡，使得每一阶段 $\pi^e = \pi = 0$。在第一阶段，雇主们持有预期 $\pi^e = 0$，在其后各阶段，如果所有前期的预期 $\pi^e = 0$，并且所有前期的实际通胀率 π 也都为 0，则持有预期 $\pi^e = 0$，否则，雇主持有预期 $\pi^e = \pi^*$。相似地，在当期预期 $\pi^e = 0$，并且所有以前的预期都为 0，所有以前的实际通胀率 π 也都为 0 时，货币当局选择令 $\pi = 0$ 的货币供给；否则，货币当局选择 $\pi = \pi^*(\pi^e)$，它是对雇主期望值 π^e 的最优反应。

假设雇主在第一阶段持有的通货膨胀预期为 $\pi^e = 0$。给定雇主们的策略（即在雇主们观测到真实的通货膨胀水平之后对其预期的调整方式），货币当局可以只集中考虑对如下两个方案的选择：① $\pi = 0$，它将使下一阶段的 $\pi^e = 0$，从而使货币当局在下一阶段仍可面临同样的选择；② 从(5 - M - 1)式计算得出的阶段最优选择 $\pi = \pi^*(0)$，它将使得此后所有的 $\pi^e = \pi^*$，这种情况下货币当局将发现此后永远选择 $\pi = \pi^*$ 也是最优的。在本期选择 $\pi = 0$，可使得每期的收益均为 $W(0,0)$，而在本期选择 $\pi = \pi^*(0)$，可使本期的收益为 $W(\pi^*(0), 0)$，但其后每期的收益永远为 $W(\pi^*, \pi^*)$。从而，当下式成立时，货币当局的选择

零通胀率的政策要优于实行通胀的政策：

$$\frac{1}{1-\delta}W(0,0) \geqslant W(\pi^*(0),0) + \frac{\delta}{1-\delta}W(\pi^*,\pi^*) \tag{5-M-2}$$

通过对上式的简化可得$\delta \geqslant c/(2c+d^2)$。参数$c$和$d$起到两方面的效果。例如，在$d$增大时，可使意料之外的通胀对产出影响的效果增大，但同样的原因，d的增大又使得阶段博弈的解π^*增大，后者又增大了货币当局的负效用（即对货币当局的惩罚）。类似地，c的增大使得通胀更加痛苦，从而使意料之外通胀的吸引力减小，但同时又使π^*下降。在两种情况下，后一作用都超过了前一影响，因而贴现因子的临界值要支持这一均衡，$c/(2c+d^2)$必须随d的增大而递减，随c的增大而递增。

至此，我们就证明了如果（5－M－2）式成立，货币当局的触发策略就是雇主触发策略的最优反应。要证明这一组策略为纳什均衡，还须证明后者是前者的最优反应。根据假设，雇主的最优策略就是正确预测通胀率，因此雇主的触发策略同样是最优的，这就证明了触发策略是无限重复货币模型的纳什均衡。证明这一组策略是子博弈完美均衡的程序与前面的分析完全一样，这里不再赘述。

㊥ 习题5.9　请你用单边背离特性证明上述三个例子（无限重复古诺双头垄断、效率工资、无通货膨胀的货币政策）中所采用的触发策略是各自子博弈的纳什均衡。

关键词

阶段博弈　重复博弈$G(T)$重复博弈的扩展式$\Gamma(T)$贴现因子δ贴现平均收益值　结果路径 v－N－M 偏好　伯努利收益函数及线性变换　重复博弈的纳什均衡　重复博弈的子博弈完美均衡　无限重复信用困境博弈　触发策略　有限惩罚策略　一报还一报策略　单边背离特性　可行收益组合　极小极大收益　无名氏定理

内容要点

1. 重复博弈是一种较特殊的动态博弈，与标准的动态博弈相比，有两个不同点：一是在博弈的每一个阶段都会产生相应的收益，而标准的动态博弈只有结束时才会产生相应的收益；二是重复博弈中参与者的偏好如果是理性的，那么必须满足 v－N－M 偏好的要求，相应的阶段博弈中的收益函数u则为伯努利收益

函数，满足线性变换。

2. 重复博弈中的纳什均衡就是要求对于整个博弈而言，参与者彼此的策略都是最优反应，而子博弈完美均衡则需考虑不同历史过程下的子博弈，有多少可能的历史就有多少可能的子博弈，只有考虑了每一种可能的子博弈，并且在该子博弈中仍是纳什均衡才能成为子博弈完美均衡。

3. 要寻找出重复博弈中的所有子博弈完美均衡是一项不可能的任务，但要证明一个策略组合是否为子博弈完美均衡却要容易得多。为了简化重复博弈中一个纳什均衡为子博弈完美均衡的证明，我们引入了单边背离特性的要求。单边背离特性是指除了子博弈的第一阶段外，一切都不变，那么即使一个参与者发生背离也不可能给其带来更大的收益。满足单边背离特性，就意味着满足子博弈完美均衡的要求，即命题 5.2。

4. 无名氏定理——可行收益组合大于纳什均衡收益。如果将纳什均衡收益组合替换为极小极大收益组合，则得到无限重复 n 人无名氏定理，但有可能不存在子博弈完美均衡。如果将博弈的参与者限定为 2 人，那么子博弈完美均衡始终是存在的，即无限重复 2 人无名氏定理。相应的，如果满足一些特定要求，那么有限重复博弈中同样存在无名氏定理——有限重复 n 人无名氏定理和有限重复 2 人无名氏定理。

进一步阅读

无名氏定理较成熟的论述见 Friedman(1971)，对无名氏定理的进一步修订与更为详细的讨论见 Abreu, Dutta and Smith(1994)和 Wen(1994)。无限重复 2 人无名氏定理可参阅 Fudenberg 和 Maskin(1991)，有限重复无名氏定理则可参阅 Benoit 和 Krishna(1985,1987)。

第六章

贝叶斯博弈

纳什均衡的关键是每一个参与者必须对其他参与者的行动持有正确的信念，而要实现这一点，参与者必须对博弈拥有足够的信息，特别是，必须知道其他参与者的偏好。然而，在多数情况下，信息是不完全的。例如，在拍卖中，竞拍者并不清楚对手对标的物的价值判断；在垄断行业，寡头彼此并不清楚对方的成本函数。所谓非完全信息是指博弈中至少有一位参与者的收益函数不是公共信息，或者说参与者的收益是不确定的。在这一章中，我们将分析非完全信息静态博弈，非完全信息博弈又被称为贝叶斯博弈（Bayesian Game）。

我们首先用一个例子来说明非完全信息博弈引起的困难，从而引出豪尔绍尼转换。第二节则主要阐述静态贝叶斯博弈的基本式和贝叶斯纳什均衡的概念。通过例子来理解所学理论仍是必不可少的，这构成了第三节的内容。在本章的最后则是显示偏好原理。

第一节　豪尔绍尼转换

我们首先来看一个例子，这个例子称为贝叶斯信用困境。

【例 6.1】　贝叶斯信用困境

假设商人有诚信和欺诈之分，讲诚信的商人被称为良商，欺诈的商人被称为奸商。为了简单，先假设商人 1 是诚信的商人，这是公共信息，但商人 2 是否讲诚信，商人 1 却并不清楚，他有可能讲诚信，也有可能从事商业欺诈。商人 2 知道自己是否讲诚信，即商人 2 是否讲诚信是一个私人信息。这个非完全信息（不对称信息）博弈可用图 6－1 表示出来。在图 6－1 中，H 表示诚信，C 表示欺诈。由于良商是讲诚信的商人，如果他欺骗一个诚信的商人，那么他会受到良心的谴责，给他带来 －1 的收益；如果商人是奸商，即使他欺骗商业伙伴也不会受到什么良心的谴责，带给他的收益为 3；如果两个商人都欺骗，那么谁也不欠谁，各得 1 的收益。

		商人2(良商)			商人2(奸商)	
		H	C		H	C
商人1	H	2,2	0,-1	H	2,2	0,3
	C	-1,0	1,1	C	-1,0	1,1

图6-1 贝叶斯信用困境博弈

在贝叶斯信用困境这个博弈中,相当于存在着两个博弈。如果商人1相信商人2是一个讲诚信的人,他们的纳什均衡为(H,H)和(C,C)[①]。如果商人1相信商人2是一个奸商,那么纳什均衡为(C,C)。但显然无法回答到底哪一个策略组合是整个博弈的纳什均衡。原因在于商人1对商人2不同的信念(belief)导致了即使是相同的策略,也会出现不同的收益,即在相同的策略组合下,收益具有不确定性。例如(H,C)这个策略组合,对参与者而言就有不同的收益,在商人2诚信的情况下,收益为(0,-1);在商人2为奸商的情况下,收益则为(0,3)。这一点显然是所有非完全信息博弈都具有的共同特性,非完全信息正是根据这一点来定义的。其次,信念不同出现的均衡解也会不同。例如,在商人1相信商人2诚信的情况下,均衡解为(H,H);而在商人1不相信商人2诚信的情况下,均衡解为(C,C)。但什么情况下相信,什么情况下不相信,并不存在标准。

由于参与者的收益函数具有不确定性,因而不可能通过求解收益最大化的方式来找到最优策略。换句话说,就是什么策略都可能成为最优策略,任何结果都有可能是博弈的均衡解。这实际上意味着我们对贝叶斯博弈没办法分析。如果博弈不是像贝叶斯信用困境那么简单,而是稍微复杂一些,情况就更是如此。

贝叶斯信用困境还可以用扩展式表示出来,如图6-2所示。为了简单,我们省掉了收益。图6-2实际上清楚地表明了贝叶斯信用困境所面临的困难,该博弈有两个开始点,在商人1行动的时候,商人1分不清他到底处于那种情况,是处于(商人2)诚信,还是处于(商人2)欺骗。由于该博弈有两个开始点,可以理解为两个不同的博弈,但关键是,这两个博弈被一条虚线连接了起来,因而它又是一个博弈。它既是两个博弈又是一个博弈,从逻辑上来说这是矛盾的,因面我们不可能直接分析它。两个开始点还意味着我们从博弈的结束点往回推,我们不知道会回到哪一个开始点。这意味着如果参与者要达到某个结束点,那么不知道采取什么样的策略才能达到这个结束点,也就是说没有人能保证一定走到这个结束点,因而均衡会不会出现只有上帝才知道了,当然最优策略也就无从谈起。

① 均衡(H, H) > (C,C),所以(H, H)是一个更为合理的纳什均衡。

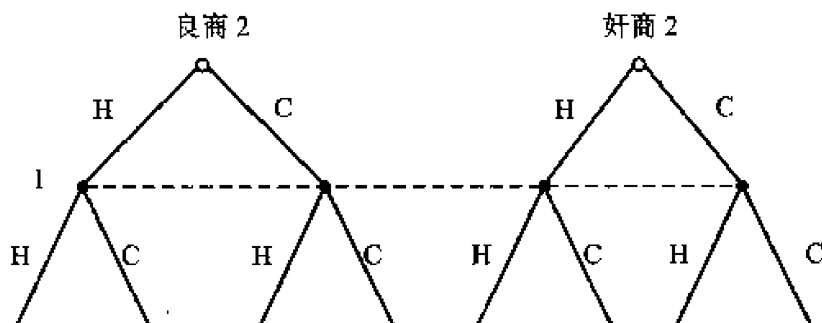


图 6-2 贝叶斯信用困境的扩展式

面对这种两个开始点的困难，一个好的想法就是引入概率分布，即商人 2 讲诚信的概率为 p，不讲诚信的概率为 $1-p$，这个概率分布由自然(Nature)决定。这样一处理，就意味着商人 1 可以计算出如果他采取某种策略，那么它可能实现的期望值。通过比较不同的期望值，就能得出作为一个"理性商人"应选择的策略，而商人 2 知道商人 1 会按照期望值的大小来决定相应的策略，因而商人 2 相应也知道他应采取什么样的策略。比如，尽管商人 2 是一个讲诚信的人，但在知道商人 1 认为他是奸商的可能性较大(也就是 p 非常小)的情况下，商人 1 选择欺诈的期望值一定会大于选择诚信的期望值，因而商人 1 的最优反应是欺诈，在这种情况下，商人 2 最优的策略只能是欺诈，尽管两个商人都是讲诚信的人。

把商人 2 是什么样的人看作一种概率分布，显然是一个非常妙的想法，它既克服了贝叶斯博弈中存在的矛盾，又能利用 v-N-M 偏好理论来分析参与者的偏好(期望值)，从而最优策略，乃至纳什均衡都有了意义。最早把这一想法变成规范性的学术语言是由经济学家豪尔绍尼在 1967 年完成的①，被称为豪尔绍尼转换。也正因为这个贡献，豪尔绍尼与纳什和塞尔腾(Selten)分享了 1994 年的诺贝尔经济学奖。

定义 6.1　豪尔绍尼转换——不同状态的出现服从一定的概率分布，并且这个概率分布是公共信息。

例如，贝叶斯信用困境中就存在两种状态——商人 2 是良商和商人 2 是奸商，而每一个状态都以一定的概率出现，不妨设每个状态出现的概率都为 1/2，转换后的贝叶斯信用困境如图 6-3 中所示，其中大方框表示商人 1 不能确定商人 2 的状态，但知道其概率各为 1/2，而小方框表示商人 2 知道自己的状态。

① Harsanyi J. "Games with Imcomplete Information Played by Bayesian Players." *Management Science*. 1967/1968, 14: 159-182, 320-334, 486-502

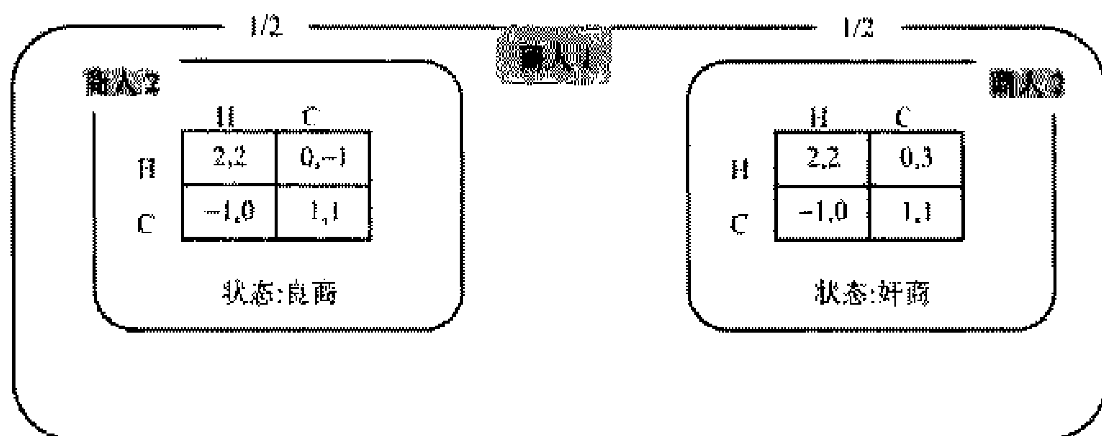


图 6-3 转换后的贝叶斯信用困境

那么所有可能策略组合下商人 1 的期望值就列在了图 6-4 中。例如,(H,(H,H))表示给定商人 2 在诚信下选择 H,在欺诈下也选择 H,那么商人 1 选择 H,其期望值将等于 2×1/2+2×1/2=2;(H,(H,C))则表示在给定商人 2 在诚信下选择 H,在欺诈下选择 C,那么商人 1 选择 H,其期望值将等于 2×1/2+0×1/2=1,其他的期望值以此类推。

		商人2			
		(H, H)	(H, C)	(C, H)	(C, C)
商人1	H	2	1	1	0
	C	-1	0	0	1

图 6-4 贝叶斯信用困境中商人 1 的期望值

这里我们可以看到,对于商人 1 而言,其策略为 H 和 C,而对于商人 2 而言,其策略则为 4 个:(H,H),(H,C),(C,H)和(C,C)。读者可能会问,商人 2 不是知道自己是什么样的人吗?为什么商人 2 的策略不是 H 和 C,却是(H,H),(H,C),(C,H)和(C,C)呢?我们可以从三个方面来理解这点:一是在博弈开始之前,商人 2 也不知道到底会出现什么状态,即不知道自己是否是诚信还是欺诈,所以他必须事先有所计划,规定好在什么样的状态下选择什么样的行动。例如商人 2 以为自己是奸商,但结果发现自己是良商,如果策略只是 C,那么商人 2 将不知道在状态为良商的情况下如何行动。二是商人 1 到市场上交易,那么有可能遇到良商,也可能遇到奸商,其概率各为 1/2。给定商人 1 选择 H,那么良商和奸商的策略显然不同,而(H,C)就表示良商选择 H,而奸商选择 C。三是虽然商人 2 知道自己是什么样的人,但商人 1 并不知道,因而商人 2 不得不把自己想象成商人 1,以便获得有关商人 1 的信念的一种正确猜测,以便选择最优行动。因而,商人 2 的策略必然由 H 和 C 所组成。

那么，贝叶斯信用困境中的纳什均衡是什么呢？根据图6－4可知，在状态的概率分布为(1/2，1/2)的情况下，如果商人2的策略为(H，C)，那么商人1的最优选择为H(因为选择C，期望值为0＜1)；反之，如果商人1选择H，那么商人2的最优策略为(H，C)(即在良商下选择诚信，收益最大为2；在奸商下选择欺诈，收益最大为3)，因而，(H，(H，C))就是贝叶斯信用困境的纳什均衡。为了增进理解，我们再举一例。

【例6.2】 罗密欧与朱丽叶

罗密欧和朱丽叶相约去听音乐会，罗密欧向朱丽叶提出邀请，但罗密欧不知道朱丽叶是否喜欢自己。如果喜欢，那么两人一起听音乐会一定要好过彼此分开；相反，如果不喜欢，那么朱丽叶就会主动回避，以免一起听音乐会遭遇尴尬。对于朱丽叶而言，也可能遇到相同的情况，如果罗密欧喜欢她，那么两人一起听音乐会要好过彼此分开；相反，如果不喜欢，那么罗密欧就会主动回避，以免尴尬。因而，该博弈就可能会出现4种状态：xx，xb，bx和bb。x代表喜欢，b代表不喜欢。例如，xb表示罗密欧喜欢朱丽叶，但朱丽叶不喜欢他。C代表古典音乐会，P代表流行音乐会。

罗密欧与朱丽叶博弈的细节都集中反映在图6－5之中，x_1和b_1代表罗密欧(参与者1)的两种状态，x_2和b_2代表朱丽叶(参与者2)的两种状态。，1：x_1代表着

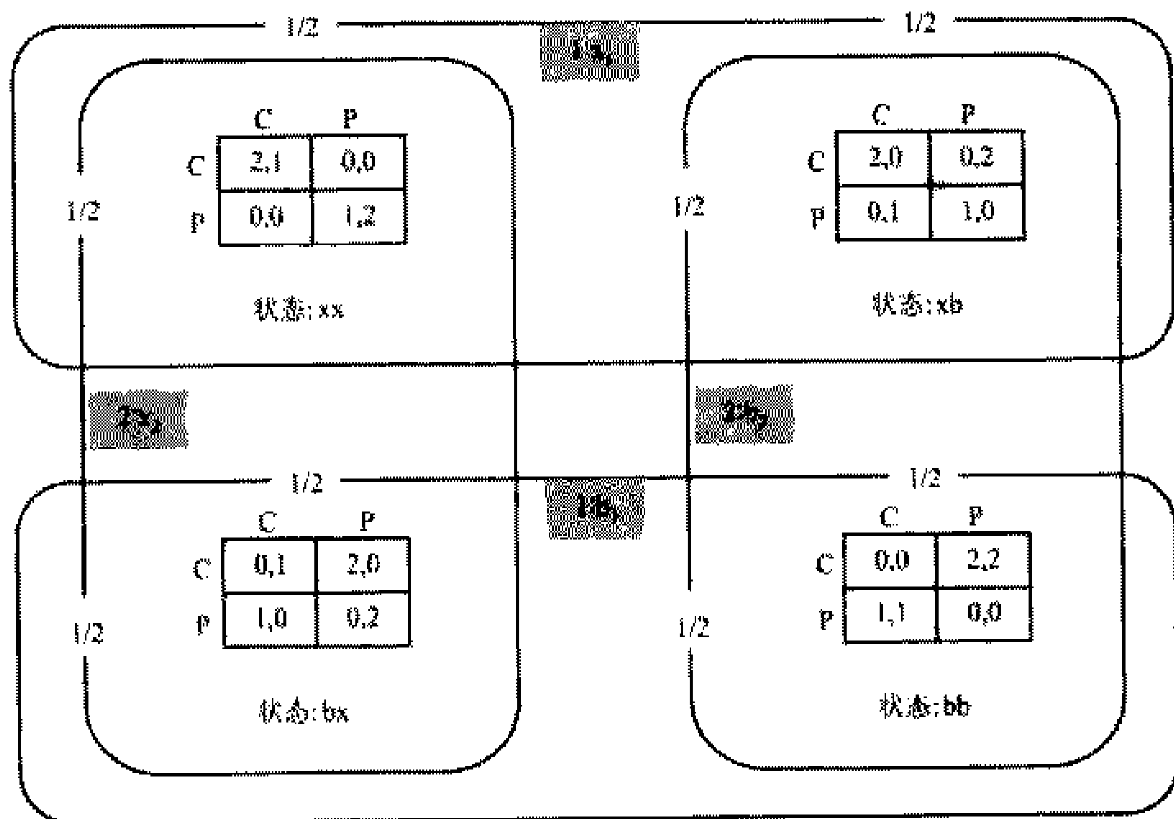


图6－5 罗密欧与朱丽叶

罗密欧的状态为喜欢朱丽叶；而 2：x_2 则代表朱丽叶的状态为喜欢罗密欧；其他类推。1/2 表示每一种状态出现的概率。对于罗密欧和朱丽叶而言，分别各有 4 个策略（C，C），（C，P），（P，C）和（P，P），运用相同的原理——即期望值，就可以得出相应的纳什均衡。

在这个博弈中，存在着两个纳什均衡：{（C，C），（C，P）} 和 {（P，C），（P，P）}。对于策略组合{（C，C），（C，P）}而言，给定朱丽叶的策略为（C，P），即状态为喜欢选择 C，状态为不喜欢选择 P，那么，如果罗密欧喜欢朱丽叶，他选择 C 的期望值为 $2\times1/2+0\times1/2=1$，而选择 P 的期望值为 $0\times1/2+1\times1/2=0.5$，因而选择 C 是罗密欧的最优反应；如果罗密欧不喜欢朱丽叶，他选择 C 的期望值为 $0\times1/2+2\times1/2=1$，而选择 P 的期望值为 $0\times1/2+1\times1/2=0.5$，因而选择 C 仍是罗密欧的最优反应。所以说，罗密欧的策略（C，C）是朱丽叶策略（C，P）的最优反应。反之，给定罗密欧的策略（C，C），那么，如果朱丽叶喜欢罗密欧，她选择 C 的期望值为 $1\times1/2+1\times1/2=1$，而选择 P 的期望值为 $0\times1/2+0\times1/2=0$，因而选择 C 是朱丽叶的最优反应；如果朱丽叶不喜欢罗密欧，她选择 C 的期望值为 $0\times1/2+0\times1/2=0$，而选择 P 的期望值为 $2\times1/2+2\times1/2=2$，因而选择 P 是朱丽叶的最优反应。所以说，朱丽叶的策略（C，P）是罗密欧策略（C，C）的最优反应。这就证明了{（C，C），（C，P）}为罗密欧与朱丽叶博弈的纳什均衡。{（C，C），（C，P）}读作：给定朱丽叶的策略（C，P），如果罗密欧喜欢朱丽叶，那么选择 C；如果不喜欢，那么选择 C。给定罗密欧的策略（C，C），如果朱丽叶喜欢罗密欧，那么选择 C；如果不喜欢，那么选择 P。

㊣ 习题 6.1 请你证明{（P，C），（P，P）}也是罗密欧与朱丽叶博弈的纳什均衡。

㊣ 习题 6.2 基于例 6.2，如果朱丽叶喜欢罗密欧的概率为 $1>q>0$，那么罗密欧与朱丽叶博弈的纳什均衡是什么？（注意 q 的取值范围）

㊥ 习题 6.3 改造贝叶斯信用困境。基于例 6.1，我们假设商人 2 同样不清楚商人 1 是否讲诚信，商人 1 讲诚信的概率为 $1>p>0$，而商人 2 讲诚信的概率为 $1>q>0$，那么改造后的贝叶斯信用博弈的纳什均衡是什么？

显然，如果将上述思想一般化，那么就能分析各种各样的贝叶斯博弈，而不仅仅是特例特办。

第二节 贝叶斯博弈与纳什均衡

在完全信息静态博弈中，参与者的策略等于行动，而在贝叶斯静态博弈中，参与者的策略是状态的函数，状态不同采取的行动也就不同。这说明贝叶斯博

弈的基本式与完全信息静态博弈的基本式有所不同,需要做一些处理。这一节中,首先给出贝叶斯博弈的基本式,然后在此基础上给出贝叶斯纳什均衡概念。

一、贝叶斯博弈的基本式

第二章中定义的 n 人完全信息静态博弈的基本式为 $G=\{N,S,u\}$,其中 $S=\{S_1,\cdots,S_n\}$,S_i为参与者 i 的策略空间;$u=(u_1,\cdots,u_n)$,$u_i(s_1,\cdots,s_n)$为所有参与者分别选择策略$(s_1,\cdots,s_n)$时参与者 i 的收益。由于完全信息静态博弈中的策略实际上等于一个简单的行动,因而基本式又可写作 $G=\{N,A,u\}$,其中 $A=\{A_1,\cdots,A_n\}$,A_i为参与者 i 的行动空间;$u=(u_1,\cdots,u_n)$,$u_i(a_1,\cdots,a_n)$为所有参与者分别选择行动$(a_1,\cdots,a_n)$时参与者 i 的收益。在贝叶斯博弈中,博弈的基本式显然不能简单地表示成这种形式,因为策略实际上依赖于参与者的状态,策略的实质是参与者根据可能的状态来选择相应的行动,而不再是简单的行动。其次,收益函数也成为状态的函数,状态不同参与者的收益函数自然不同,如果收益函数与参与者的状态无关,就意味着贝叶斯博弈中的收益函数退化为完全信息静态博弈中的收益函数,完全信息静态博弈就成为贝叶斯博弈的特例。所谓状态就是对参与者特性集合的一个完整描述,包含着参与者的偏好和相关信息。不同的状态意味着不同的特性集合,所有可能的状态就构成了状态集合。用 z 来表示某个状态,$Z=\{z_1,z_2,\cdots\}$就表示状态集合。

在博弈开始时,某个状态成为现实,但(有些)参与者并不能直接观察到状态(否则就是完全信息了),而是观察到某种**信号**,这个信号可能反映出有关状态的某些信息。我们用 $\tau_i(z)$来表示参与者 i 在状态 z 发生时观察到的信号,函数 τ_i 就称为参与者 i 的**信号函数**。显然,如果对于每一个参与者而言,任意 $z\neq y$ 意味着 $\tau_i(z)\neq\tau_i(y)$,那么就不存在不确定性,因而如果信息是不完全的,就一定存在着某些状态发出的信号一样(这样参与者才分不清到底是哪种状态)。凡是给出信号 t_i的状态的集合就称为与信号 t_i 相一致的**状态类**①,简称状态类 t_i,数学上可表示为$\{z:\tau_i(z)=t_i\}$。如果每一个参与者的每一个状态类只包含一个状态,就说明这个博弈是一个完全信息博弈,另一个极端情况则是只存在一个状态类(即所有状态发出的信号都一样),因而参与者 i 状态类的大小实际上反映了参与者 i 拥有信息的充分性。例如,存在三个状态 z_1,z_2和 z_3,$\tau_i(z_1)=t_1$,$\tau_i(z_2)=\tau_i(z_3)=t_2$,因而当参与者 i 观察到信号为 t_1时,参与者 i 知道发生的是状态 z_1,但如果观察到 t_2,参与者 i 只知道发生了状态 z_2或 z_3,但不能确定到底是

① 有些教材是从参与者的类型来考虑不确定性的,实际上参与者的类型就是参与者接受到的信号。

哪一个状态。参与者 i 的信念就是定义在状态类上的概率分布。例如，状态类 t_1 只包含一个状态，因而其概率就是 1，而状态类 t_2 包含两个状态，因而有关状态 z_2 和 z_3 的信念就是一个概率分布 $(p,1-p)$。在非完全信息静态博弈中，参与者的偏好是定义在行动和状态之上的，即收益函数为 $u(a,z)$，其中 a 为行动组合。由于参与者对状态不确定，因而参与者的偏好还必须考虑与向量 (a,z) 相应的概率分布，因而参与者的收益函数必须为伯努利收益函数，即满足 v－M－N 理性偏好的要求。

实际上，对于参与者的信念，准确的定义应该考虑条件概率，而不是简单概率。我们用 $Pr(z\mid t_i)$ 来表示参与者 i 在观察到信号 t_i 的条件下，对状态 z 可能发生概率的一个猜测（信念），并且参与者会根据贝叶斯法则不断对这种猜测进行修正。通常为了分析简单，我们假设各状态之间彼此独立。

有了上面这些说明，现在就可以给出贝叶斯博弈的正式定义。

定义 6.2　贝叶斯博弈的基本式表示为 $G_B=\{N,Z,A,T,P,u\}$，其中 N 为参与者集合；Z 为状态集合；A 为行动集合；T 为信号集合；P 则为概率分布集合，它实际上代表着参与者的信念，对于参与者 i 而言，状态 z 的后验概率为 $Pr(z\mid t_i)$；$u=(u_1,\cdots,u_n)$，u_i 是参与者 i 定义在 (a,z) 上的伯努利收益函数，因而参与者 i 的期望收益 $v_i=Pr(z_1\mid t)u_i(a,z_1)+\cdots+Pr(z_k\mid t)u_i(a,z_k)$，其中状态类 $t=\{z_1,z_2,\cdots,z_k\}$。

由于策略是一个从状态到行动的映射，通过行动和状态可以构建出所有的策略，所以参与者的策略空间没有在贝叶斯博弈的基本式 G_B 中出现。

对于参与者的信念 P，有几点需要重点说明。

第一，概率 $Pr(z)$ 由自然决定，并成为公共信息，但另一方面它又可以理解成是参与者的一种先验概率。这种先验概率不是凭空产生，而是来自于（理性的）参与者对自然所有的了解下最科学地得出的。它既是客观的，又是主观的。比如，中央电视台每晚的天气预报，它是利用最科学的手段（即充分理性）根据所有能获得的（自然）信息，对明天下雨的概率作出的主观猜测，由于它是现有技术条件下所能得到的天气情况最好的预测，所以它又是客观的，并被所有的电视观众所知（公共信息）。通常为了简化分析，一般假设各状态彼此独立。

第二，概率 $Pr(z\mid t_i)$ 表示参与者 i 在观察到信号 t_i 的前提下，对状态 z 发生的概率的信念。它实际上是参与者 i 根据贝叶斯法则（Bayes Rule）计算得到的后验概率。对于其他参与者而言，他们知道参与者 i 对 z 所持有的推断，反之也一样，即所有 $Pr(z\mid t_i)$ 是一个公共信息。

如果 z 是离散的，参与者 i 观察到信号 t，那么

$$Pr(z \mid t) = \frac{Pr(z)}{\sum_{z \in \{z\}_t} Pr(z)} \quad (6-1)$$

其中$\{z\}_t$表示状态类t，即$\{z: \tau_i(z)=t\}$。

第三，贝叶斯博弈G_B实际上是由若干个策略式博弈G组成，只不过每一个不同的状态对应着一个不同的策略式博弈。这实际上意味着我们求贝叶斯博弈只不过是在计算一串策略式博弈的平均值，豪尔绍尼转换的意义就在于此。

例6.1 贝叶斯信用困境的基本式$G_B=\{N,Z,A,T,P,u\}$，其中：

(1) 参与者集合：$N=\{$商人1，商人2$\}$。

(2) 状态集合：$Z=\{$良商，奸商$\}$。

(3) 行动集合：$A_i=\{H,C\}$，$i \in N$。

(4) 信号集合：$T=\{\{t_1\},\{t_{21},t_{22}\}\}$，对于商人1，$\tau_1$(良商)$=\tau_1$(奸商)$=t_1$，但对于商人2，$\tau_2$(良商)$=t_{21}$，$\tau_2$(奸商)$=t_{22}$。

(5) 信念：对于商人1，Pr(良商$\mid t_1$)$=Pr$(奸商$\mid t_1$)$=1/2$，而对于商人2，Pr(良商$\mid t_{21}$)$=1$，Pr(奸商$\mid t_{21}$)$=0$，或者Pr(奸商$\mid t_{22}$)$=1$，Pr(良商$\mid t_{22}$)$=0$。

(6) 偏好：对于商人1而言，存在两个$u_1(a;$良商$)$和$u_1(a;$奸商$)$，期望收益就为$v_1=1/2\times u_1(a;$良商$)+1/2\times u_1(a;$奸商$)$，而对于商人2，其期望收益为$v_2=1\times u_2(a;$良商$)+0\times u_2(a;$奸商$)$或$v_2=0\times u_2(a;$良商$)+1\times u_2(a;$奸商$)$。

例6.2 罗密欧与朱丽叶的基本式$G_B=\{N,Z,A,T,P,u\}$，其中：

(1) 参与者集合：$N=\{$罗密欧，朱丽叶$\}$。

(2) 状态集合：$Z=\{xx,xb,bx,bb\}$。

(3) 行动集合：$A_i=\{C,Pop\}$，$i\in N$。（为了防止与概率分布P相混，用Pop来表示流行乐）

(4) 信号集合：$T=\{\{x_1,b_1\},\{x_2,b_2\}\}$，即$\tau_1(xx)=\tau_1(xb)=x_1$，$\tau_1(bx)=\tau_1(bb)=b_1$，$\tau_2(xx)=\tau_2(bx)=x_2$，$\tau_2(xb)=\tau_2(bb)=b_2$。

(5) 信念：对于罗密欧，$Pr(xx\mid x_1)=Pr(xb\mid x_1)=1/2$，$Pr(bx\mid b_1)=Pr(bb\mid b_1)=1/2$；对于朱丽叶，$Pr(xx\mid x_2)=Pr(bx\mid x_2)=1/2$，$Pr(xb\mid b_2)=Pr(bb\mid b_2)=1/2$。

(6) 偏好：在喜欢的状态下，罗密欧的期望收益为$v_1=1/2\times u_1(a,xx)+1/2\times u_i(a,xb)$；在不喜欢的状态下，$v_1=1/2\times u_1(a,bx)+1/2\times u_i(a,bb)$。朱丽叶的期望收益类似，不再赘述。

在罗密欧与朱丽叶博弈中，根据已知条件可知每个状态发生的概率都为1/4，（例如罗密欧喜欢朱丽叶的概率为1/2，而朱丽叶喜欢罗密欧的概率也为

1/2,所以两人彼此喜欢发生的概率为 $1/2 \times 1/2 = 1/4$。)而根据贝叶斯法则,可知 $Pr(\mathrm{xx} \mid \mathrm{x}_1) = Pr(\mathrm{xx})/(Pr(\mathrm{xx}) + Pr(\mathrm{xb})) = 1/2$。

【例 6.3】 贝叶斯 3 人古诺模型

该模型所有的内容和第二章提到的内容一样,不一样的地方在于:① 参与者为三个垄断企业,定义为 1,2,3;② 第 i 个企业的成本有可能为高成本 c_H^i 和低成本 c_L^i,$i=1,2,3$。第一个企业为低成本的概率为 θ_1,高成本的概率为 $1-\theta_1$;第二个企业、第三个企业依此类推,即 θ_2 和 $(1-\theta_2)$,θ_3 和 $(1-\theta_3)$。每个企业知道自己的成本,但不能确定其他企业的成本,因而存在八种状态,即 $\{c_L^1c_L^2c_L^3, c_L^1c_L^2c_H^3, \cdots, c_H^1c_H^2c_H^3\}$。为了简单起见,假设各状态彼此独立。

贝叶斯古诺模型的基本式 $G_B = \{N, Z, A, T, P, u\}$,其中:

(1) 参与者集合:$N = \{1,2,3\}$。

(2) 状态集合:$Z = \{\{c_L^1c_L^2c_L^3, c_L^1c_L^2c_H^3, c_L^1c_H^2c_L^3, c_L^1c_H^2c_H^3, c_H^1c_L^2c_L^3, c_H^1c_L^2c_H^3, c_H^1c_H^2c_L^3, c_H^1c_H^2c_H^3\}$。

(3) 行动集合:产量 q_i 就是企业 i 的行动,行动空间 $A_i = [0, a]$。

(4) 信号集合:$T = \{c_L^1, c_H^1, c_L^2, c_H^2, c_L^3, c_H^3\}$,对于企业 1 而言,$\tau_1(c_L^1c_L^2c_L^3) = \tau_1(c_L^1c_L^2c_H^3) = \tau_1(c_L^1c_H^2c_L^3) = \tau_1(c_L^1c_H^2c_H^3) = c_L^1$,$\tau_1(c_H^1c_L^2c_L^3) = \tau_1(c_H^1c_L^2c_H^3) = \tau_1(c_H^1c_H^2c_L^3) = \tau_1(c_H^1c_H^2c_H^3) = c_H^1$,即企业 1 知道自己的成本,但不知道其他企业的成本。其他类推。

(5) 信念:对于企业 1 而言,在观察到信号 c_L^1 后,对企业 2 和企业 3 都出现低成本的概率判断为

$$Pr(c_L^2c_L^3 \mid c_L^1) = \frac{Pr(c_L^1c_L^2c_L^3)}{Pr(c_L^1)} = \theta_2\theta_3$$

其他类推。

(6) 偏好:企业 1 在收到信号 c_L^1 后的期望收益函数为

$v_1 = \theta_2\theta_3u_1(a, c_L^1c_L^2c_L^3) + \theta_2(1-\theta_3)u_1(a, c_L^1c_L^2c_H^3) + (1-\theta_2)\theta_3u_1(a, c_L^1c_H^2c_L^3) + (1-\theta_2)(1-\theta_3)u_1(a, c_L^1c_H^2c_H^3)$。

其他类推。

㊀ 习题 6.4 请给出例 6.3 中贝叶斯博弈基本式中的所有信念和偏好。

通过上面的说明,读者现在应该清楚为什么会把非完全信息博弈称为贝叶斯博弈,因为参与者信念的形成是根据贝叶斯法则来完成的。另外,为了简化分析,我们定义参与者的类型就是参与者接受到的信号,例如,企业 1 可能接收到的信号为 c_L^1 和 c_H^1,因而企业 1 的类型就有两个 c_L^1(低成本)和 c_H^1(高成本)。

二、贝叶斯纳什均衡

实际上,豪尔绍尼转换的意义就在于将非完全信息处理成一种完全信息,从而使得我们可以像处理完全信息博弈一样来处理贝叶斯博弈,这样纳什均衡就同样适合于贝叶斯博弈。不同之处在于参与者的偏好要满足 v-M-N 理性偏好的要求,即期望收益函数。在此情况下,贝叶斯博弈的参与者的策略就是指出在什么样的信号下采取什么样的行动,否则就不可能求出参与者的期望收益。

定义 6.3 贝叶斯博弈 $G_B=\{N,Z,A,T,P,u\}$,参与者 i 的一个**纯策略** s_i是一个由 T_i映射到 A_i的函数。对于 T_i中的每一个 t_i,$s_i(t_i)$表明当参与者 i 观察到信号 t_i时,参与者 i 将从行动空间 A_i中选择一个行动。

由于 T_i既是信号集合,又是参与者 i 的类型集合,所以策略也可以理解为关于类型的函数,有时为了方便我们也这样使用。

例如,在贝叶斯信用困境中,商人 1 只观察到一个信号,因而策略就是行动,策略空间 = {H,C};而商人 2 却观察到两个信号,因而策略就是一个行动向量,策略空间 = {(H,H),(H,C),(C,H),(C,C)}。

㊂ 习题 6.5 请你总结一下策略式博弈、完全信息动态博弈和贝叶斯博弈中有关策略的定义,并谈一谈你的理解。

给出了贝叶斯博弈关于策略的定义之后,我们就可以定义贝叶斯纳什均衡。

定义 6.4 贝叶斯博弈 $G_B=\{N,Z,A,T,P,u\}$,策略组合 $s^*=(s_1^*,\cdots,s_n^*)$ 是一个**纯策略贝叶斯纳什均衡**,如果对每一个参与者 i,及对 i 的信号空间 T_i中的每一个信号 t_i,$s_i^*(t_i)$满足

$$\sum_{z\in\{z:\tau_i(z)=t_i\}} Pr(z\mid t_i)u_i[s_1^*(\tau_1(z)),\cdots,s_i^*(t_i),\cdots s_n^*(\tau_n(z))]\geqslant$$
$$\sum_{z\in\{z:\tau_i(z)=t_i\}} Pr(z\mid t_i)u_i[s_1^*(\tau_1(z)),\cdots,s_i(t_i),\cdots s_n^*(\tau_n(z))] \quad (6-2)$$

其中,$s_i^*(t_i)\neq s_i(t_i)$。需要注意的是 $\tau(z)=t$,$s(t)=a$,所以 $u(a,z)$可以紧凑地写作 $u(s(t))$。

㊂ 习题 6.6 请你运用公式(6-2)验证贝叶斯信用困境和罗密欧与朱丽叶博弈的纳什均衡。

式(6-2)虽然有许多的数学符号,但中心思想却非常简单,与完全信息静态博弈中的纳什均衡并没有什么两样;每一个参与者的策略必须是其他参与者策略的最优反应,并且对每一个参与者都如此。

在前面我们曾提到,贝叶斯博弈 G_B实际上是由若干个策略式博弈 G 组成,这里我们相应地给出与贝叶斯博弈和贝叶斯均衡完全等价的策略式博弈和纳什

均衡。

命题 6.1 贝叶斯博弈 G_B 的贝叶斯纳什均衡与如下定义的策略式博弈 G 的纳什均衡完全等价。对于策略式博弈 $G=\{N,A,u\}$，其中：

(1) N 为向量 (i,t_i) 的集合，每一个 (i,t_i) 都代表一个参与者。

(2) A 为行动空间，$A(i,t_i)$ 就是每一个参与者 (i,t_i) 的行动集合。

(3) 偏好：每一个参与者 (i,t_i) 的伯努利收益函数为 $\sum Pr(z\mid t_i)u_i[(a_i, a_{-i}(z));t_i]$，其中 $a_{-i}(z)$ 表示其他参与者在状态为 z 的情况下采取的行动。

如果读者对于理解贝叶斯博弈觉得有困难，但却能够很好地理解完全信息静态博弈，那么就可以运用命题 6.1 来求贝叶斯博弈的均衡解。实际上，求贝叶斯博弈只不过是在计算若干个策略式博弈的平均值，只不过权重为概率 $Pr(z\mid t_i)$。

例如，贝叶斯信用困境的策略式博弈 $G=\{N,A,u\}$ 中：

(1) 参与者集合：$N=\{(1,\mathrm{H}),(2,\mathrm{H}),(2,\mathrm{C})\}$。(1,H) 为诚信商人 1，(2,H) 为诚信商人 2，(2,C) 代表欺诈商人 2。

(2) 行动空间：$A_{1,\mathrm{H}}=\{\mathrm{H},\mathrm{C}\}$，$A_{2,\mathrm{H}}=\{\mathrm{H},\mathrm{C}\}$，$A_{2,\mathrm{C}}=\{\mathrm{H},\mathrm{C}\}$。

(3) 伯努利收益函数：对于商人 1，期望收益函数为

$$\frac{1}{2}u_1(a_1,a_{2,\mathrm{H}};t_1)+\frac{1}{2}u_1(a_1,a_{2,\mathrm{C}};t_1)$$

对于商人 (2,H)，其收益函数为

$$u_{2,\mathrm{H}}(a_1,a_{2,\mathrm{H}};t_{21})\qquad (\text{因为 } Pr(\mathrm{H}\mid t_{21})=1)$$

对于商人 (2,C)，其收益函数为

$$u_{2,\mathrm{C}}(a_1,a_{2,\mathrm{C}};t_{22})\qquad (\text{因为 } Pr(\mathrm{C}\mid t_{22})=1)$$

其中，$t_1=\mathrm{H}$，表示商人 1 收到的信号，$t_{21}=\mathrm{H}$ 和 $t_{22}=\mathrm{C}$，表示商人 2 收到的信号。根据纳什均衡的定义，很容易得到 (H,(H,C)) 为贝叶斯信用困境的纳什均衡（即商人 1 与商人 (2,H) 的纳什均衡为 (H,H)，商人 1 与商人 (2,C) 的纳什均衡为 (H,C)，记住商人 1 的收益是期望收益），它也是该博弈的贝叶斯纳什均衡。

下面我们给出混合策略贝叶斯纳什均衡，跳过并不影响后面的阅读。

混合策略的贝叶斯纳什均衡

定义 6.5 在贝叶斯博弈 $G_B=\{N,Z,A,T,P,u\}$ 中，参与者 i 的一个混合策略 s_i 是一个由 T_i 映射到 ΔA_i 的函数。ΔA_i 为定义在行动空间 A_i 上的概率分布空间。$s_i(t_i)$ 代表当参与者 i 的信号为 t_i 时，他所采用的混合策略。对于任意一个行动 $a_i\in A_i$，$s_i(t_i)(a_i)$ 表示在混合策略 $s_i(t_i)$ 下，a_i 被使用的概率。$s=(s_1, s_2,\cdots,s_n)$ 为一个混合策略组合，ΔS_i 为参与者 i 所有可能混合策略的集合，而

$\Delta S=(\Delta S_1,\cdots,\Delta S_n)$是所有可能策略组合的集合。

如参与者i的信号是t_i,则参与者i在混合策略组合s下的收益为

$$\tilde{u}_i(s;t_i)=\sum_{z\in\{z:T_i(z)=t_i\}}Pr(z\mid t_i)\sum_{(a_1,\cdots a_n)\in A}s_1(t_1)(a_1)\times\cdots\times s_n(t_n)(a_n)u_i[(a_1,\cdots,a_n),z]$$

定义6.6 贝叶斯博弈$G_B=\{N,Z,A,P,T,u\}$,策略组合$s^*=(s_1^*,\cdots,s_n^*)$是一个**混合策略贝叶斯纳什均衡**,如果对每一个参与者i,及对i的信号空间T_i中的每一个信号t_i,$s_i^*(t_i)$满足

$$\tilde{u}_i(s_i^*,s_{-i}^*;t_i)\geqslant\tilde{u}_i(s_i,s_{-i}^*;t_i)$$

对所有的$s_i\in\Delta S_i$成立。

定理6.1 贝叶斯纳什均衡存在性定理。对于有限的贝叶斯静态博弈$G_B=\{N,Z,A,T,P,u\}$,即参与者是有限的,A是有限的,Z是有限的,那么一定存在(混合)贝叶斯纳什均衡。

定理的证明与纳什均衡的证明一样。通过豪尔绍尼转换,我们知道贝叶斯静态博弈可以处理成一串基本式博弈的加权平均——期望值。而每一个基本式博弈都存在着纳什均衡(纳什定理),因而n个基本式博弈线性相加(即加权平均)之后的博弈一定存在着纳什均衡。因为线性相加不会改变任何函数的拓扑性质,所以每一个基本式博弈存在纳什均衡,就意味着贝叶斯静态博弈存在纳什均衡——即贝叶斯纳什均衡。实际上,在满足上述定理条件的情况下,在动态博弈中,只要博弈的次数不是无穷的,那么动态博弈中也一定存在纳什均衡。

作为一个练习,罗密欧与朱丽叶博弈的混合策略下的收益是什么呢?设罗密欧在状态x_1下的行动C的概率为p_{x_1},在状态b_1下的行动C的概率为p_{b_1},朱丽叶相应的概率为q_{x_2}和q_{b_2}。根据已知条件,对于x_1的罗密欧,如果给定朱丽叶也采用混合策略,那么需要考虑两种情况,一是朱丽叶喜欢罗密欧,二是朱丽叶不喜欢罗密欧。在朱丽叶喜欢罗密欧的情况下,罗密欧选择行动C的收益为

$$q_{x_2}u_1(\mathrm{C,C;xx})+(1-q_{x_2})u_1(\mathrm{C,P;xx}) \qquad ①$$

在朱丽叶不喜欢罗密欧的情况下,罗密欧选择行动C的收益为

$$q_{b_2}u_1(\mathrm{C,C;xb})+(1-q_{b_2})u_1(\mathrm{C,P;xb}) \qquad ②$$

因而罗密欧选择行动C的预期收益为

$$1/2\times① + 1/2\times② \qquad ③$$

在朱丽叶喜欢罗密欧的情况下,罗密欧选择行动P的收益为

$$q_{x_2}u_1(\mathrm{P,C;xx})+(1-q_{x_2})u_1(\mathrm{P,P;xx}) \qquad ④$$

在朱丽叶不喜欢罗密欧的情况下，罗密欧选择行动 P 的收益为

$$q_{b_2}u_1(P,C;xb)+(1-q_{b_2})u_1(P,P;xb)$$ ⑤

因而罗密欧选择行动 P 的预期收益为

$$1/2\times④+1/2\times⑤$$ ⑤

如果要罗密欧选择混合策略，那么式③就必须等于式⑤，由此得出的结论是：当 $q_{x_2}+q_{b_2}=2/3$ 时，x_1 的罗密欧将选择混合策略。

对于 b_1 的罗密欧，同样需要考虑两种情况。给定罗密欧选择 C，在朱丽叶喜欢罗密欧的情况下，其收益为

$$q_{x_2}u_1(C,C;bx)+(1-q_{x_2})u_1(C,P;bx)$$ ⑦

在朱丽叶不喜欢罗密欧的情况下，其收益为

$$q_{b_2}u_1(C,C;bb)+(1-q_{b_2})u_1(C,P;bb)$$ ⑧

因而罗密欧选择 C 的预期收益为

$$1/2\times⑦+1/2\times⑧$$ ⑨

给定罗密欧选择 P，那么在朱丽叶喜欢罗密欧的情况下，其收益为

$$q_{x_2}u_1(P,C;bx)+(1-q_{x_2})u_1(P,P;bx)$$ ⑩

在朱丽叶不喜欢罗密欧的情况下，其收益为

$$q_{b_2}u_1(P,C;bb)+(1-q_{b_2})u_1(P,P;bb)$$ ⑪

因而罗密欧选择 P 的预期收益为

$$1/2\times⑩+1/2\times⑪$$ ⑫

如果罗密欧选择混合策略，那么式⑨就必须等于式⑫，由此得出的结果为：当 $q_{x_2}+q_{b_2}=4/3$ 时，b_1 的罗密欧选择混合策略。

上述的分析总结如下：对于 x_1 的罗密欧，当 $q_{x_2}+q_{b_2}=2/3$ 时，罗密欧选择混合策略；对于 b_1 的罗密欧，当 $q_{x_2}+q_{b_2}=4/3$ 时，罗密欧选择混合策略。

同理，可以得出朱丽叶选择混合策略的条件。对于 x_2 的朱丽叶，给定朱丽叶选择 C，在罗密欧喜欢朱丽叶的情况下，其收益为

$$p_{x_1}u_2(C,C;xx)+(1-p_{x_1})u_1(P,C;xx)$$ ⑬

在罗密欧不喜欢朱丽叶的情况下，其收益为

$$p_{b_1}u_2(C,C;bx)+(1-p_{b_1})u_1(P,C;bx)$$ ⑭

因而朱丽叶选择 C 的预期收益为

$$1/2\times⑬+1/2\times⑭$$ ⑮

给定朱丽叶选择 P，其预期收益为

$$1/2\times[p_{x_1}u_2(C,P;xx)+(1-p_{x_1})u_2(P,P;xx)+$$
$$1/2\times[p_{b_1}u_2(C,P;bx)+(1-p_{b_1})u_2(P,P;bx)]$$ ⑯

如果朱丽叶选择混合策略，那么式⑮必须等于式⑯，由此得出的结果为：当 $p_{x_1}+p_{b_1}=4/3$ 时，x_2的朱丽叶会选择混合策略。利用相同的方法，可以得出：对于 b_2的朱丽叶，当 $p_{x_1}+p_{b_1}=2/3$ 时，朱丽叶会选择混合策略（这里不再赘述，请读者自行证明）。

㊁ 习题 6.7　请你验证 b_2的朱丽叶选择混合策略的条件为 $p_{x_1}+p_{b_1}=2/3$，同时请证明罗密欧与朱丽叶博弈是否存在混合的贝叶斯纳什均衡。

第三节　经典举例

下面将通过一些经典例子来巩固前面所学知识。

一、贝叶斯 2 人古诺模型

该模型所有内容与贝叶斯 3 人古诺模型一样，只是为了简单和便于读者理解，将企业数设定为 2，同时假设 $c_L^1=c_L^2=c_L$，$c_H^1=c_H^2=c_H$。其基本式省略，读者可以根据贝叶斯 3 人古诺模型照葫芦画瓢（你可以把它作为一个练习）。根据已知条件，得

$$Pr\ (c_L^2\mid c_L^1)\pi_1(q_1,q_2(c_L^2);c_L^1)\ =\ \theta_2 q_1(a-q_1-q_2(c_L^2)-c_L^1) \qquad ①$$

$$Pr\ (c_H^2\mid c_L^1)\pi_1(q_1,q_2(c_L^2);c_L^1)\ =\ (1-\theta_2)q_1(a-q_1-q_2(c_H^2)-c_L^1) \qquad ②$$

企业 1 在观察到信号 c_L后的期望收益函数就是期望利润函数，即

$$E\pi_1(q_1,q_2;c_L)\ =\ ①+②$$

根据纳什均衡的定义，应使利润最大化，即

$$\max_{q_1}\quad E\pi_1(q_1,q_2;c_L^1)$$

求一阶条件，并整理得

$$q_1(c_L)\ =\ \frac{a-[\theta_2 q_2(c_L)+(1-\theta_2)q_2(c_H)]-c_L}{2} \qquad (6-G-1)$$

同理，企业 1 观察到信号 c_H后的产量为

$$q_1(c_H)\ =\ \frac{a-[\theta_2 q_2(c_L)+(1-\theta_2)q_2(c_H)]-c_H}{2} \qquad (6-G-2)$$

利用对称性可以相应得到企业 2 的策略：

$$q_2(c_L)\ =\ \frac{a-[\theta_1 q_1(c_L)+(1-\theta_1)q_1(c_H)]-c_L}{2} \qquad (6-G-3)$$

$$q_2(c_H)\ =\ \frac{a-[\theta_1 q_1(c_L)+(1-\theta_1)q_1(c_H)]-c_H}{2} \qquad (6-G-4)$$

策略组合$[(q_1^*(c_L),q_1^*(c_H)),(q_2^*(c_L),q_2^*(c_H))]$如果是(6-G-1)、

(6-G-2)、(6-G-3)、(6-G-4)式的解,那么它就是贝叶斯纳什均衡。

㊐ 习题6.8 如果上例的 $\theta_1=\theta_2=1/2$,那么请你解出具体的均衡值。

㊐ 习题6.9 请你给出贝叶斯3人古诺模型的贝叶斯纳什均衡。

二、贝叶斯公共产品

在一个由 n 人组成的社区中存在着一种公共产品,只要有一位社区成员愿意承担提供公共产品的成本,那么公共产品就能提供,否则公共产品将会消失。不妨设提供公共产品的成本为 $c>0$,如果公共产品能够提供,那么公共产品对社员 i 的价值为 v_i;如果公共产品不能提供,那么收益为0。每一个社员都不知道其他人的 v,但知道 v 是位于 v_{min} 和 v_{max} 之间的一个连续随机变量,其累积分布函数为 $F(v)$,并且每个社员的价值 v 彼此独立。所有的社区成员同时选择某个行动,为了简单不妨设行动要么为 c(表示愿意承担提供公共产品的成本),要么为0(表示不愿意承担提供公共产品的成本)。$v_{max}>c>v_{min}>0$。只要有一个社员选择了 c,那么公共产品就能被提供。对于社员 i 而言,当其行动为 c 时,其收益函数为 v_i-c;当其他社员行动为 c,而他的行动为0时,收益函数为 v_i;当所有社员行动为0时,收益函数为0。

贝叶斯公共产品博弈的基本式 $G_B=\{N,Z,A,T,P,u\}$,其中:

(1) 参与者集合:n 个社员组成,$N=\{1,\cdots,n\}$。

(2) 状态集合:$Z=\{z\}$,$z=(v_1,\cdots,v_n)$,即所有社员价值判断组合的集合。

(3) 行动集合:$A_i=\{0,c\}$,$i\in N$。

(4) 信号集合(类型集合):$\tau_i(z)=v_i$,$v_{max}\geq v_i\geq v_{min}$。

(5) 信念:给定社员 i 的类型 v_i,$Pr(z\mid v_i)=F(v_1)F(v_2)\cdots F(v_{i-1})F(v_{i+1})\cdots F(v_n)$,因为各 v 是独立的。

(6) 收益函数:社员 i 在状态 z 下的伯努利收益函数为

$$\begin{cases}v_i-c & a_i=c\\ v_i & a_j=c,a_i=0,j\neq i\\ 0 & (a_1,\cdots,a_n)=0\end{cases}$$

如果信息是完全的,那么很容易证明当 $v_i\geq c$,那么社员 i 提供公共产品,而其他社员不提供是一个纳什均衡。即使信息不完全,这同样是一个贝叶斯纳什均衡,只不过相对应的是其他类型的社员 i 和所有类型的其他社员。

㊐ 习题6.10 请证明 $v_i\geq c$,类型为 v_i 的社员 i 提供公共产品,而其他类型的社员 i 和所有类型的其他社员不提供公共产品是一个纳什均衡。

显然如果对于所有社员,如果 $c\geq v$,那么唯一的均衡就是没有人愿意提供公共产品,比如国防和治安,其成本如此之大,不可能由某个社会成员独自承担,而

需要通过一种机制以便将成本在社会成员之间进行分摊。因而,一个国家的税制和社会保障制度可以运用博弈论来进行合理的设计,这属于机制设计理论研究的内容。

那么在该例子中,是否还存在着其他的均衡?这里不妨考虑一个对称贝叶斯纳什均衡,其中每个社员的策略为:存在一个临界值 v^*,如果社员 i 的 $v_i \geq v^*$,那么选择 c;如果 $v_i < v^*$,那么选择 0。给定其他社员的策略如此,那么需要证明的就是对于任一社员 i 在 $v_i < v^*$ 时,选择 0 的期望值要大于选择 c 的期望值;在 $v_i \geq v^*$ 时,选择 c 的期望值要大于选择 0 的期望值。

在 $v_i < v^*$ 时,要求选择 0 的期望值要大于选择 c 的期望值,即 Pr(至少有一个社员选择 $c \mid v_i$) $\times v_i + [1 - Pr$(至少一个社员选择 $c \mid v_i$)$] \times 0 > v_i - c$(选择 c 意味着公共产品肯定提供)。其中 Pr(至少一个社员选择 $c \mid v_i$) $= 1 - Pr$(没有一个社员选择 $c \mid v_i$),根据已知条件可知,一个社员不选择 c 的概率为 $F(v^*)$,即所有 $v < v^*$ 的概率。由于每个社员的 v 是独立的,因而 Pr(没有一个社员选择 $c \mid v_i$) $= F(v^*)^{n-1}$。所以选择 0 的期望值要大于选择 c 的期望值就要求

$$[1 - F(v^*)^{n-1}]v_i \geq v_i - c \qquad (6-P-1)$$

如果 $v_i \geq v^*$,那么选择 c 的期望值就要大于选择 0 的期望值,即

$$[1 - F(v^*)^{n-1}]v_i \leq v_i - c \qquad (6-P-2)$$

要同时满足(6-P-1)与(6-P-2)式,并且对每一个社员 i 都成立,只有

$$v^* F(v^*)^{n-1} = c \qquad (6-P-3)$$

即如果 v^* 满足上式,那么上面提到的策略就是最优策略,这就证明存在一个对称的贝叶斯纳什均衡。由于 $v_{max} > c$,因而存在 $v^* > c$,这表明在该均衡中存在 $v^* > v_i \geq c$ 却无法提供公共产品的可能。实际上,只要 $v_i \geq c$,那么社员 i 提供公共产品将会使得整个社会的福利增加,但这种帕累托改进却无法实现。

一个有趣的问题是,如果社员的人数增加,那么提供公共产品的可能性是增还是减?根据(6-P-3)式,当 n 增加时,v^* 必须增加才能保证(6-P-3)式成立。没有一个社员选择 c 的概率为 $F(v^*)^n$,因而至少有一个社员选择 c 的概率就为 $1 - F(v^*)^n$,同样根据(6-P-3)式,有 $F(v^*)^n = cF(v^*)/v^*$,因而提供公共产品的概率就等于

$$1 - cF(v^*)/v^*$$

如果 $F(v^*)/v^*$ 是 v^* 的增函数,那么提供公共产品的概率就是 v^* 的减函数;如果 $F(v^*)/v^*$ 是 v^* 的减函数,那么提供公共产品的概率就是 v^* 的增函数。不妨假设 $F(v^*)$ 是一个均匀分布,那么 $F(v^*)/v^*$ 就是关于 v^* 的增函数,这就意味着如果 n 增加,那么提供公共产品的概率会下降,即

$$n\uparrow \rightarrow v^*\uparrow \rightarrow F(v^*)/v^*\uparrow \rightarrow 1-F(v^*)/v^*\downarrow$$

实际上,人越多表明搭便车的情况越多,因而提供公共产品的可能性变小,也许这就是奥尔森所描述的"集体行动的逻辑"。

三、陪审团

在欧美国家,对嫌疑人是否有罪的判定是由陪审团作出的。陪审团成员根据呈堂证供进行表决,只有陪审团成员全体一致地认为嫌疑人有罪,嫌疑人才被判有罪,否则将无罪释放。每一个陪审团成员如何理解证据显然是一个私人信息。在决定如何投票上,陪审团成员必然会考虑将一个无罪之人判为有罪或将有罪之人判为无罪的可能,同时考虑在给定其他成员的投票时,其投票可能对审判结果产生的影响。因而我们可以把陪审团成员看作是一个贝叶斯博弈中的参与者,通过对该博弈的分析,来揭示有罪之人被判无罪与无罪之人被判有罪的可能性。

不妨假设嫌疑人有罪的先验概率为 π,并且为一个公共信息。嫌疑人有罪或无罪是一种状态,但陪审团成员无法直接观察到嫌疑人的状态,而只能观察到信号——证据。在看到呈堂证供后,陪审团成员会对相关的概率进行修正,为了简单,不妨设每一个成员的概率都相等。在嫌疑人有罪的情况下,陪审团成员根据证据认为嫌疑人有罪的概率为 p;在嫌疑人无罪的情况下,陪审团成员根据证据认为嫌疑人无罪的概率为 q。为了表明证据是有用的,假设 $p>1/2$ 和 $q>1/2$,因而 $p>1-q$。每一个陪审团成员都希望将罪犯送人监狱,而将无罪之人释放,两种结果无差异,并严格好于将无辜者判人狱或将有罪者释放,因而每一个陪审团成员的伯努利收益函数设为

$$u=\begin{cases}0 & \text{如果有罪人狱或无罪释放}\\ -z & \text{如果无罪人狱}\\ -(1-z) & \text{如果有罪释放}\end{cases}$$

我们如此设置收益函数的原因在于 z 实际上是陪审团成员认定嫌疑人有罪的概率的临界值。我们不妨设想一个陪审团成员是如此投票的:由于状态不可能直接观察到,因而陪审团成员的一个自然想法就是根据一个概率临界值来决定如何投票,比如,存在一个 z,当 Pr(嫌疑人有罪|给定陪审团成员的类型)$>z$,陪审团成员投票有罪;如果 Pr(嫌疑人无罪|给定陪审团成员的类型)$\leqslant z$,那么陪审团成员投票无罪。设 r 为陪审团成员在所有能用的信息下认定嫌疑人有罪的概率,那么如果嫌疑人被释放,陪审团成员的预期收益为 $-r(1-z)+(1-r)0=-r(1-z)$;如果嫌疑人被判有罪,其预期收益为 $r\times 0-(1-r)z=-z(1-r)$。因而当 $-r(1-z)>-(1-r)z$(即 $z>r$)时,陪审团成员投无罪的票

的期望收益才会大于投有罪的票的期望收益，因而陪审团成员会投无罪票；当 $z<r$ 时，则会投有罪票。这表明，z 实际上是陪审团成员认定嫌疑人有罪的概率临界值，这就是为什么在参与者的收益函数中会出现 z 的原因。

每一个陪审团成员的行动为 C(投票有罪)或 Q(投票无罪)，嫌疑人的真实状态为 G(有罪)或 I(无罪)，陪审团成员 i 接收到的信号为 s_i，它表示 i 对证据的看法，其可能为 g(指向有罪)或 b(指向无罪)，$(s_1,\cdots,s_n)$ 则为 n 个陪审团成员的信号向量。因而一个状态就是向量 $(X,s_1,\cdots,s_n)$，其中 $X=\{G,I\}$。在状态 $X=G$ 的情况下，每一个陪审团成员认为其他参与者接到信号 g 的概率为 p，而接收到信号 b 的概率为 $1-p$。在状态 $X=I$ 的情况下，每一个陪审团成员认为其他参与者接到信号 g 的概率为 $1-q$，而接受到信号 b 的概率为 q。所有参与者接收到信号的概率彼此独立。

那么陪审团博弈的基本式为 $G_B=\{N,Z,A,T,P,u\}$，其中：

(1) 参与者集合：n 个陪审团成员。

(2) 状态集合：$Z=\{z\}$，$z=(X,s_1,\cdots,s_k)$，其中 $X=\{G,I\}$，$s_i=\{g,b\}$。$s_i=g$ 表示参与者 i 接收到的信号为有罪，$s_i=b$ 表示参与者 i 接收到的信号为无罪。

(3) 行动集合：每一个参与者 i 的行动空间 $A_i=\{C,Q\}$，C 表示投有罪的票，Q 表示投无罪的票。

(4) 信号集合：每一个参与者 i 接收到的信号集合 $T_i=\{g,b\}$，相应的信号函数为 $\tau_i(X,s_1,\cdots,s_k)=s_i$，即每一个参与者只知道自己接收到的信号，而对嫌疑人的状态和其他参与者的信号并不知情。

(5) 信念：对于类型为 g 的陪审团成员 i，其认为状态 $z=(G,s_1,\cdots,s_k)$ 的概率为 $Pr((G,s_1,\cdots,s_k)\mid g)=\pi p^{k-1}(1-p)^{n-k}$；认为状态为 $z=(I,s_1,\cdots,s_k)$ 的概率为 $Pr((I,s_1,\cdots,s_k)\mid g)=(1-\pi)(1-q)^{k-1}q^{n-k}$。对于类型为 b 的陪审团成员 i，其认为状态 $z=(G,s_1,\cdots,s_k)$ 的概率为 $Pr((G,s_1,\cdots,s_k)\mid b)=\pi p^k(1-p)^{n-k-1}$；认为状态为 $z=(I,s_1,\cdots,s_k)$ 的概率为 $Pr((I,s_1,\cdots,s_k)\mid b)=(1-\pi)(1-q)^k q^{n-k-1}$。上面的 k 表示有 k 个参与者收到的信号为 g。其他以此类推。

(6) 收益函数：陪审团成员 i 的伯努利收益函数为

$$u_i(a,z)=\begin{cases}0 & a\neq(C,\cdots,C)\text{ 且 }X=I\text{ 或 }a=(C,\cdots,C)\text{ 且 }X=g\\ -z & a=(C,\cdots,C)\text{ 和 }X=I\\ -(1-z) & a\neq(C,\cdots,C)\text{ 和 }X=G\end{cases}$$

陪审团博弈存在着许多的策略，要直接寻找陪审团博弈的纳什均衡将会十分困难。我们不妨考虑这样一个策略：即如果收到的信号为 b，就投票无罪；如果收到的信号为 g，就投票有罪。我们不妨把它称为“直觉策略”。

我们由易入难，分别考虑三种情况：① $n=1$；② $n=2$；③ $n>2$。

第一种情况：只有一位陪审团成员，这等价于由法官作出判决。

如果法官接受到的信号为 g（即证据指向有罪），法官认定其有罪的概率为

$$Pr(G \mid g)=\frac{Pr(g \mid G)Pr(G)}{Pr(g \mid G)Pr(G)+Pr(g \mid I)Pr(I)}=\frac{p\pi}{p\pi+(1-q)(1-\pi)} \quad (6-J-1)$$

因而当且仅当 $Pr(G \mid g)>z$，法官才会判决有罪，否则判决无罪。

如果法官接收到的信号为 b（即证据指向无罪），法官认定其有罪的概率为

$$Pr(G \mid b)=\frac{Pr(b \mid G)Pr(G)}{Pr(b \mid G)Pr(G)+Pr(b \mid I)Pr(I)}=\frac{(1-p)\pi}{(1-p)\pi+q(1-\pi)} \quad (6-J-2)$$

那么当且仅当 $Pr(G \mid b)<z$，法官才会判决无罪，否则判决有罪。因而，如果直觉策略为最优策略就必须要求 z 同时满足(6－J－1)式和(6－J－2)式，由于 $p>1/2, q>1/2$，容易证明，(6－J－1)＞(6－J－2)，因而存在 z 使得下式成立。

$$\frac{(1-p)\pi}{(1-p)\pi+q(1-\pi)}<z<\frac{p\pi}{p\pi+(1-q)(1-\pi)} \quad (6-J-3)$$

即如果临界值 z 满足(6－J－3)不等式，那么直觉策略就为最优策略。

第二种情况：存在两位陪审团成员。

在参与者 2 采用直觉策略的情况下，参与者 1 采用直觉策略是否为最优反应呢？考虑参与者 1 接收到的信号为 g，那么参与者 2 有可能接收到信号 g 或 b，如果接收到的信号为 b，那么无论参与者 1 如何做，嫌疑人都将被释放（因为只有全体一致投票有罪才能判嫌疑人入狱），这说明参与者 1 可以忽略不计参与者 2 的类型为 b 的情况，而只需考虑参与者 2 的类型为 g 的情况。在参与者 2 接收到的信号为 g 的情况下，如果直觉策略为参与者 1 的最优策略，那么参与者 1 接收到 g，并投票 C 的期望值应该大于投票 Q 的期望值。在参与者收到的信号都为 g 的情况下，参与者 1 认定嫌疑人有罪的概率为

$$Pr(G \mid g,g)=\frac{Pr(g,g \mid G)Pr(G)}{Pr(g,g \mid G)Pr(G)+Pr(g,g \mid I)Pr(I)}=\frac{p^2\pi}{p^2\pi+(1-q)^2(1-\pi)} \quad (6-J-4)$$

要让参与者 1 投票有罪，必须满足式(6－J－4)＞z。

同理，如果参与者 1 接收到的信号为 b，那么参与者 1 会忽略参与者 2 的类

型为 b 的情况(因为参与者 1 无论如何做,嫌疑人都将被释放),因而只需要考虑参与者 2 的类型为 g 的情况。那么在参与者 1 的类型为 b,参与者 2 的类型为 g,参与者 1 认定嫌疑人有罪的概率为

$$Pr(G \mid b, g) = \frac{Pr(b,g \mid G)Pr(G)}{Pr(b,g \mid G)Pr(G) + Pr(b,g \mid I)Pr(I)}$$

$$= \frac{(1-p)p\pi}{(1-p)p\pi + (1-q)q(1-\pi)} \quad (6-J-5)$$

因而如果要让参与者 1 投票无罪,就必须满足式(6-J-5) $< z$。

综上所述,给定参与者 2 的直觉策略,参与者 1 的直觉策略为最优反应,当且仅当存在 z 同时满足(6-J-4)和(6-J-5)式,即下面的不等式成立。

$$\frac{(1-p)p\pi}{(1-p)p\pi + (1-q)q(1-\pi)} < z < \frac{p^2\pi}{p^2\pi + (1-q)^2(1-\pi)} \quad (6-J-6)$$

利用对称性可知,给定参与者 1 的直觉策略,参与者 2 的直觉策略为最优反应。因而直觉策略就是 2 人陪审团博弈的贝叶斯纳什均衡。

(易) 习题 6.11　请你比较式(6-J-3)和式(6-J-6)有什么区别,并作出解释。

第三种情况:$n > 2$ 陪审团。

仍然考虑直觉策略。不妨设参与者 1 接收到的信号为 g,与两人的情况一样,如果 $n-1$ 个参与者中至少有一人接收到的信号为 b,那么无论参与者 1 如何做,结果都一样,因而参与者 1 可以忽略这种情况,而集中精力考虑其他所有参与者接受到的信号都为 g 的情况。要使得投票有罪的期望值大于投票无罪的期望值,等同于要求认定嫌疑人有罪的概率必须大于 z。在所有信号为 g 的情况下,参与者 1 认定嫌疑人有罪的概率为

$$Pr(G \mid g, \cdots, g) = \frac{Pr(g,\cdots,g \mid G)Pr(G)}{Pr(g,\cdots,g \mid G)Pr(G) + Pr(g,\cdots,g \mid I)Pr(I)}$$

$$= \frac{p^n\pi}{p^n\pi + (1-q)^n(1-\pi)} \quad (6-J-7)$$

因而如果参与者 1 投票有罪,必须满足式(6-J-7) $> z$。

同理,若参与者 1 接收到的信号为 b,那么参与者 1 会忽略其他参与者类型为 b 的情况,而集中精力考虑所有其他参与者类型为 g 的情况。在参与者 1 的类型为 b、其他参与者的类型为 g 的条件下,参与者 1 认定嫌疑人有罪的概率为

$$Pr(G \mid b, \cdots, g) = \frac{Pr(b,\cdots,g \mid G)Pr(G)}{Pr(b,\cdots,g \mid G)Pr(G) + Pr(b,\cdots,g \mid I)Pr(I)}$$

$$=\frac{(1-p)p^{n-1}\pi}{(1-p)p^{n-1}\pi+q(1-q)^{n-1}(1-\pi)} \qquad (6-J-8)$$

因而如果参与者 1 投票无罪,必须满足式$(6-J-8)<z$。

综上所述,给定其他参与者的直觉策略,参与者 1 的直觉策略为最优反应,当且仅当存在 z 同时满足$(6-J-7)$和$(6-J-8)$式,即下面的不等式成立。

$$\frac{(1-p)p^{n-1}\pi}{(1-p)p^{n-1}\pi+q(1-q)^{n-1}(1-\pi)}<z<\frac{p^n\pi}{p^n\pi+(1-q)^n(1-\pi)} \qquad (6-J-9)$$

利用对称性可知,上述分析对于任一参与者 i 都是成立的。因而直觉策略就是 n 人陪审团博弈的贝叶斯纳什均衡。

但上述分析是建立在 n 不能很大的基础之上的,因为当 n 越大(即陪审团成员越多),式$(6-J-8)$将趋于 1。因为 $p>1-q$,当 $n\to\infty$,$p^{n-1}/(1-q)^{n-1}\to\infty$。这意味着陪审团成员根据直觉策略来投票有可能($n\to\infty$)不是贝叶斯纳什均衡,因为 $1<z<1$ 显然是矛盾的。如果存在纳什均衡,无论陪审团成员接受到的信号是什么,嫌疑人被判有罪的可能性将非常高。例如,如果 $p=q=0.8$,$\pi=0.5$,而 $n=7$,那么式$(6-J-8)>0.999$,即如果陪审团成员根据直觉策略来投票,那么嫌疑人基本上 100% 将被判有罪。

需要指出的是,上述分析没有任何的价值判断意义,不能就此认为由法官判决就要好于陪审团制度,因为模型是建立在高度抽象和直觉策略之上的。实际上,陪审团博弈里面存在许多的其他策略,有些策略可能更为合理和有效。

㊊ 习题 6.12　如果$(6-J-8)\leqslant z\leqslant(6-J-7)$,那么对于 n 人陪审团博弈,直觉策略是否是贝叶斯纳什均衡?

㊊ 习题 6.13　设陪审团成员数为任意整数,请你给出使得所有陪审团成员不管其信号是什么都投票有罪或投票无罪为贝叶斯纳什均衡的条件。

实际上,该博弈还存在着一个对称混合策略纳什均衡:即类型为 g 的参与者以 1 的概率投票有罪,类型为 b 的参与者以一定的概率投票有罪和无罪。定义$(m,1-m)$为类型 b 的参与者的混合策略,即投票有罪的概率为 m,投票无罪的概率为 $1-m$。与前面的分析相同,参与者 1 只需要考虑其他参与者投票为 C 的情况(因为如果有人投票为 Q,那么无论参与者 1 如何投票,嫌疑人都将被释放)。那么在嫌疑人有罪的前提下,其他参与者投票 C 的概率为 $p+(1-p)m$(即信号为 g 投票 C,信号为 b 则以 m 的概率投票 C);在嫌疑人无罪的前提下,其他参与者投票 C 的概率为$(1-q)+qm$。不妨设参与者 1 的类型为 g,那么在给定其他参与者投票 C 的情况下,参与者 1 认定嫌疑人有罪的概率为

$$Pr(G \mid g, C, \cdots, C) = \frac{Pr(g \mid G)Pr(C \mid G)^{n-1}Pr(G)}{Pr(g \mid G)Pr(C \mid G)^{n-1}Pr(G) + Pr(G \mid I)Pr(C \mid I)^{n-1}Pr(I)}$$

$$= \frac{p[p + (1-p)m]^{n-1}\pi}{p[p + (1-p)m]^{n-1}\pi + (1-q)(1-q+qm)^{n-1}(1-\pi)} \tag{6-J-10}$$

如果参与者 1 投票 C,就要求(6 - J - 10)式$\geqslant z$。

同理,如果参与者 1 的类型为 b,那么参与者 1 认定嫌疑人有罪的概率为

$$Pr(G \mid b, C, \cdots, C) = \frac{Pr(b \mid G)Pr(C \mid G)^{n-1}Pr(G)}{Pr(b \mid G)Pr(C \mid G)^{n-1}Pr(G) + Pr(b \mid I)Pr(C \mid I)^{n-1}Pr(I)}$$

$$= \frac{(1-p)[p + (1-p)m]^{n-1}\pi}{(1-p)[p + (1-p)m]^{n-1}\pi + q(1-q+qm)^{n-1}(1-\pi)} \tag{6-J-11}$$

参与者 1 随机进行投票要求(6 - J - 11)式 $= z$。显然,$z =$ (6 - J - 11) $\leqslant$ (6 - J - 10),因而只要 $z =$ (6 - J - 11)式,给定其他参与者选择这种混合策略,那么这种混合策略同样是参与者 1 的最优策略。利用对称性,可知这种混合策略组合为陪审团博弈的一个混合策略贝叶斯纳什均衡。

上面的分析是在全体一致原则的基础上进行的,如果不要求全体一致,比如少数服从多数(这个分析可能很复杂),并且陪审团成员都采取类似于上面提到的混合策略,那么随着 n 的增加,一个无辜者被判有罪的概率将趋于 0(见 Feddersen and Pesendorfer, 1998 的论文),这说明陪审团制度在防止冤假错案上可能是非常有效的(但会使得一些罪犯漏网,不过根据犯罪心理学,罪犯侥幸逃脱后通常会再次作案,因而罪犯被判刑的可能性是趋于 1 的,即所谓"法网恢恢,疏而不漏")。由于欧美国家奉行无罪推定的原则,因而陪审团制度在欧美国家就较为流行。

四、拍卖理论

贝叶斯博弈最重要的一个运用就是对拍卖机制的分析。最理想的拍卖机制应该能够将竞拍人的真实偏好显示出来,而参与者真实偏好的显示在经济学中具有重要的意义。因为公共产品供给、外在性解决、公司治理、产权配置等问题,核心就在于参与者有动机"虚假"地显示其偏好。也正因为如此,对拍卖机制的分析直接成为信息经济学研究的发端,对拍卖机制作出开创性研究的维克瑞(Vickery)也因此获得了 1996 年诺贝尔经济学奖。

依据不同的方式,拍卖大致可分为以下几种类型:

(1) 英式拍卖(English auction)。拍卖时拍卖师通常会报出一个起始价格,

然后由竞拍人开始报价,后一个报价必须比前一个报价高,当最后一个竞拍人报出的价格高到没有人愿意应拍时,拍卖品就以该价成交。英式拍卖在拍卖所是最常见的一种拍卖形式,中央电视台黄金时段的广告就是以这种方式竞拍的。

(2) 荷式拍卖 (Dutch auction)。这种拍卖方式与英式拍卖正好相反,因最早在荷兰的拍卖所运用而得名。拍卖师首先会报出一个非常高的价格,如果没有人应拍,竞拍师再报一个较低的价格,如果还没有人应拍,再向下报价,一直到有人应拍为止,拍卖品就以这个价格成交。

(3) 一级密封式价格拍卖(sealed - bid first price auction)。顾名思义,就是所有的竞拍者的报价只有竞拍者本人和拍卖者知道,其他竞拍者不知道,只有报价最高的竞拍者才能中标。例如,在公共工程招标中,所有竞标者将自己的叫价放在密封的信封中交给发标人,然后发标人公布中标者。

(4) 二级密封式价格拍卖(sealed - bid second price auction)。这和密封式一级价格拍卖的方式一样,也使用保密的方式来竞标。唯一不同的是报价最高的竞拍者中标,但出的价格只是第二高价者的价格。这种拍卖方式是维克瑞发明的,又称维克瑞拍卖。维克瑞 1962 年证明,这种拍卖机制能够诱使竞拍者显示他的真实偏好,作为诱使竞拍者显示真实偏好的代价就是拍卖者只能得到第二高价,第一高价与第二高价的差价就是对竞拍者说真话的奖励。这实际上涉及激励兼容的问题,激励兼容问题是整个信息经济学的基石。

对于某个标的物,每一个竞拍者独立地观察到有关标的物价值的某种信息——即信号,如果这个信号就是其对标的物的价值判断,那么我们说竞拍者的价值是私人的。私人价值意味着竞拍者的信号彼此独立,互不影响,因而价值也互不影响。相反,如果每一个竞拍者对标的物的价值判断不仅取决于他所接受到的信号,而且还取决于其他竞拍者的信号,那么竞拍者的价值就是公共的。公共价值意味着竞拍者的信号彼此不独立,竞拍者之间对拍卖品的价值评估就具有相关性。

严格来讲,英式拍卖和荷式拍卖是一个动态博弈,而密封式拍卖为静态博弈。但从本质上看,在私人价值和风险中性下,英式拍卖的预期收益以密封式二级价格拍卖的预期收益为极限,竞拍者在荷式拍卖与密封式一级价格拍卖中所面临的情况则完全一样。为什么?因为英式拍卖中第一高价竞拍者可以等到其他竞拍者不再出价时出手,同时第一高报价越接近第二高报价将使第一高竞拍者的预期收益越多(所以只要比第二高报价高一点点就可以了),这一点与二级价格拍卖一样。因为在二级价格拍卖中,每一个竞拍者报价等于自己的价值评估为优策略,在优策略均衡下,第一高竞拍者支付的价格只是第二高价。对于荷式拍卖与密封式一级价格拍卖,竞拍者所面临的情况则完全一样,因为荷式拍卖

的叫价从高到低，因而第一高竞拍者只有叫价低于其对标的物的价值判断时，才可能出手。另一方面，其越等待（即叫价越低），预期收益越多，但失去标的物的可能性也越大。在密封式一级价格拍卖中，只有报价低于标的物的价值判断才不是劣策略（报价大于等于竞拍者的价值判断是劣策略）。同样，报价越低，收益越多，但失去标的物的可能性越大。因而对于竞拍者而言，荷式拍卖与密封式一级价格拍卖并无二致。读者可能会猜想，如果密封式一级价格拍卖与二级价格拍卖的预期收益相等，那么上述四种拍卖制度不就是一样的吗？这就是所谓的等收益定理。

㊂ 习题 6.14 请你证明在密封式一级价格拍卖中，竞拍者报价 $b_i(v_i)$ 大于等于其对标的物的价值判断 v_i 相对于 $b_i(v_i) < v_i$ 是一个劣策略。（为了简单，你也可以假设信息完全。）

命题 6.1 等收益定理（revenue equivalent theorem）在私人价值拍卖中，竞拍者 i 对标的物的价值判断为 v_i，$i=1,2,\cdots,n$，且每一个 v_i 都服从区间为 $[v_{\min},v_{\max}]$ 的随机分布 $F(v)$，$F(v)$ 为累积分布函数；每个竞拍者都为风险中性者；则拍卖者或每一个竞拍者在上述四种拍卖制度下的预期收益（在极限上）都是一样的①。

下面我们将分别在私人价值与公共价值的情况下，考察密封式一级价格拍卖与二级价格拍卖。

（一）独立私人价值

每一个竞拍者对标的物的价值判断都是私人信息，虽然不知道其他竞拍者的价值判断，但所有价值判断都服从区间为 $[v_{\min},v_{\max}]$ 的累积分布函数 $F(v)$。如果没有中标，那么收益为 0；如果中标，那么收益为 $(v-p)/m$，p 为实际支付价格，而 m 则为赢得标的物的人数。那么，私人价值密封式一级价格拍卖与二级价格拍卖的贝叶斯博弈为 $G_B=\{N,Z,A,T,P,u\}$，其中：

(1) 参与者集合：$N=\{1,2,\cdots,n\}$，即存在 n 个竞拍者。

(2) 状态集合：$Z=\{z\}$，$z=(v_1,v_2,\cdots,v_n)$，即所有竞拍者的价值向量构成一个状态，其中 $v_{\min}\leqslant v_i\leqslant v_{\max}$，$i\in N$。

(3) 行动集合：每一个竞拍者 i 的行动就是给出一个报价 $b_i\geqslant 0$。

(4) 信号集合：对于每一个竞拍者 i 而言，其接收到的信号只是自己的价值判断，并不知道其他竞拍者的价值，即 $\tau_i(v_1,\cdots,v_i,\cdots,v_n)=v_i$。

① 更一般的分析见 Myerson R. B. "Optimal Auction Design", *Mathematics of Operations Research*. 1981, 6: 58－73

(5) 信念:对于每一个竞拍者 i 而言,给定其类型 v_i,其他每一个竞拍者 j 的价值不超过 v_j 的概率为 $F(v_1)\times\cdots\times F(v_{i-1})\times F(v_{i+1})\times\cdots\times F(v_n)$。

(6) 偏好:每一个竞拍者 i 的伯努利收益函数为

$$u_i(b_i,(v_1,\cdots,v_n)) = \begin{cases}[v_i - p(b_i)]/m & b_i \geqslant b_j \\ 0 & b_i < b_j\end{cases}$$

其中 b 表示报价,m 表示包括 i 在内有 m 个竞拍者的报价最高。如果 $p(b_i)=b_i$,上述贝叶斯博弈即为一级价格拍卖;如果 $p(b_i)=b_j$,b_j 为第二高价,上述博弈则为二级价格拍卖。

1. 二级价格拍卖的贝叶斯纳什均衡

在完全信息静态博弈中,我们曾经分析过二级价格拍卖,得出的结论是:每一个竞拍者的报价等于其价值,即 $b(v)=v$ 是一个优策略(见例 2.6)。因而每一个竞拍者都报出其价值就构成一个纳什均衡。那么,在贝叶斯博弈中,每一个竞拍者都报出其价值是否也是一个贝叶斯纳什均衡呢?答案是肯定。

给定其他竞拍者都报出其价值,那么竞拍者 i 报出其价值是否是最优呢?

第一,$b(v_i)=v_i$ 要优于 $b(v_i)>v_i$。如果其他竞拍者 j 的价值 v_j 都小于 v_i,那么 $b(v_i)=v_i$ 与 $b(v_i)>v_i$ 都能得到标的物,收益一样,因而 $b(v_i)=v_i$ 与 $b(v_i)>v_i$ 一样好。如果某个竞拍者 j 的价值 v_j 大于 v_i,那么 $b(v_i)>v_i$ 有可能中标(也可能不中标),这时支付的价格为 $v_j>v_i$,收益小于 0,不如报 $b(v_i)=v_i$,不中标,收益等于 0,因而 $b(v_i)=v_i$ 要优于 $b(v_i)>v_i$。如果存在某些竞拍者的价值 v_j 等于 v_i,那么 $b(v_i)>v_i$ 如果中标,其收益为 0,但如果其报价为 $b(v_i)=v_i$,如果中标,收益为 $(v_i-p)/m\geqslant 0$,如果不中标收益都为 0,因而 $b(v_i)=v_i$ 同样优于 $b(v_i)>v_i$。

第二,$b(v_i)=v_i$ 要优于 $b(v_i)<v_i$。如果其他竞拍者 j 的价值 v_j 都小于 v_i,那么 $b(v_i)=v_i$ 一定能获得标的物,收益为正。如果报 $b(v_i)<v_i$,那么可能中标,与 $b(v_i)=v_i$ 收益一样,也可能不中标,其收益为 0,显然 $b(v_i)=v_i$ 优于 $b(v_i)<v$。如果某个竞拍者 j 的价值 v_j 大于 v_i,那么 $b(v_i)=v_i$ 与 $b(v_i)<v_i$ 都不可能中标,收益都为 0,因而 $b(v_i)=v_i$ 与 $b(v_i)<v_i$ 一样好。如果存在某些竞拍者的价值 v_j 等于 v_i,那么 $b(v_i)<v_i$ 肯定不能中标,其收益为 0,但如果其报价为 $b(v_i)=v_i$,那么有可能中标,收益为 $(v_i-p)/m\geqslant 0$,如果不中标收益为 0,因而 $b(v_i)=v_i$ 同样优于 $b(v_i)<v_i$。

综上所述,对于所有类型的其他竞拍者 j,无论竞拍者 i 的类型如何,其报价等于价值是优策略,因而对于所有类型的竞拍者,其报价等于价值就构成一个贝叶斯纳什均衡。这表明,在二级价格拍卖中,每一个竞拍者都报出自己的价值是一个优策略,这一点与博弈中的信息是否完全无关。同时也说明,维克瑞拍卖具有良好的稳定性,并不因信息的多寡而变。二级价格拍卖还存在许许多多的贝

叶斯均衡，但只有报价等于价值是唯一的优策略组成的均衡。

㊣ 习题6.15　二级价格拍卖的缺陷。如果竞拍者 i 公开宣称无论付出多大的代价，对拍卖品志在必得，请证明在二级价格拍卖中，对于任一竞拍者 i，存在一个贝叶斯纳什均衡，其结果为竞拍者 i 赢得标的物。（提示：你可以考虑信息完全的情况）

2. 一级价格拍卖的贝叶斯纳什均衡

在一级价格拍卖中，每一个竞拍者报价大于等于其价值，即 $b(v)\geqslant v$，都要劣于报价小于其价值，即 $b(v)<v$。原因非常简单，如果 $b(v)\geqslant v$，那么无论竞拍者的类型是什么，也无论其他竞拍者如何，其收益一定小于0（如果中标），最好的情况是等于0（如果不中标）；相反，如果报价 $b(v)<v$，那么其收益要么等于0（如果不中标），要么大于0（如果中标）。所以，在一级价格拍卖中，竞拍者只可能选择小于其价值的报价，这一点与信息的多寡无关，这也是二级价格拍卖的价值所在。此外，一般而言，$b(v)$ 为增函数而不可能为减函数（即价值越高，报价越高），如果 $b(v)$ 为减函数，就意味着价值越高，报价越低，唯一的可能只能是价值最低的竞拍者获得拍卖品。这不仅与常理相悖，而且也违背了拍卖的宗旨。

我们首先分析 $n=2$ 的情况，然后再考虑 $n>2$ 的情况。

为了便于读者理解，我们假定 $v_{min}=0, v_{max}=1, F(v)$ 为均匀分布。即使竞拍者只有2人，要找到一级价格拍卖的贝叶斯均衡也是困难的，我们不妨先猜测每一个竞拍者的报价为其价值的一半是一个贝叶斯均衡，即 $b_i(v_i)=v_i/2, i=1,2$。

不失一般性，给定竞拍者2的报价为 $b_2(v_2)=v_2/2$，那么 b_2 为位于0到1/2的均匀分布。要赢得拍卖品，除非竞拍者1的报价超过 b_2，其概率为

$$Pr\{b_1(v_1)>b_2(v_2)\mid v_1\}=Pr\{b_1(v_1)>v_2/2\}\quad（信号彼此独立）$$

因而，当 $b_1>1/2$ 时，他肯定能够获得拍卖品；当 $b_1<1/2$ 时，他获得拍卖品的概率为 $Pr\{b_1>b_2\}=2b_1$（因为 b_2 为 $[0,1/2]$ 上的均匀分布），由此可得竞拍者1的收益函数为

$$\begin{cases}2b_1(v_1-b_1) & 0\leqslant b_1\leqslant 1/2\\ v_1-b_1 & b_1>1/2\end{cases}$$

即竞拍者1的预期收益函数为分段函数，容易证明当 $b_1(v_1)=v_1/2$ 时，竞拍者1的预期收益最大。竞拍者1的预期收益函数如图6－6所示。从图6－6可以看出，当 $b_1=v_1/2$ 时（假设 $v_1<1$），预期收益达到最大值。利用对称性，相同的结论对竞拍者2同样成立。这就证明了每一个竞拍者的报价 $b(v)=v/2$ 为一个贝叶斯纳什均衡。

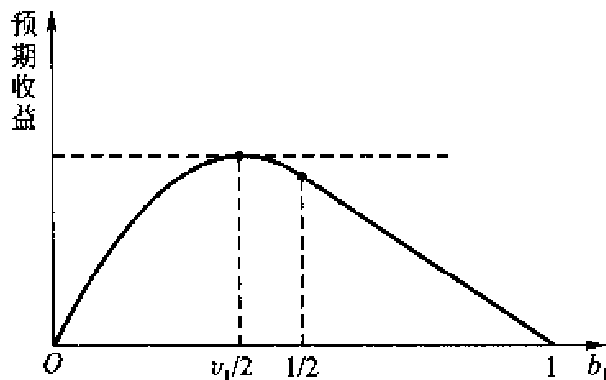


图 6-6 竞拍者 1 的预期收益函数

在 $n>2$ 的情景下，同样存在相似的结论，即存在这样一个贝叶斯纳什均衡，其中每一个竞拍者的报价为其价值的 $1-1/n$，即 $b_i(v)=(1-1/n)v$。为了更为一般化，我们放弃 $F(v)$ 为区间 $[0,1]$ 上的均匀分布这样的假设，$F(v)$ 可以为任意分布。下面的分析会运用到微分和概率分布的一些知识，但都是最基本的，由于贝叶斯博弈涉及随机性，完全不用数学将很难将问题说清，因而读者应该有勇气克服它。

报价函数仍满足前面提到的条件，但假设所有竞拍者的报价函数都一样，即 $b_i(v_i)=b(v_i)$，不妨称为公共报价函数。由于报价函数都一样，我们把相应的均衡称为对称均衡。不失一般性，对于任一竞拍者，如果其报价 b 并获得拍卖品，那么其收益为

$$(v-b)Pr\{\text{最高报价为 } b\} = (v-b)Pr\{\text{其他报价} \leqslant b\}\text{①} \qquad (6-S-1)$$

由于 $b(v)$ 为单调增函数，因而存在反函数 $b^{-1}(b)=v$。一个竞拍者如果其报价不超过 b，那么其价值也不会超过 $b^{-1}(b)$，这就意味着报价不超过 b 的概率为 $F(b^{-1}(b))$，因而 $n-1$ 个竞拍者报价都不超过 b 的概率就为 $(F(b^{-1}(b)))^{n-1}$。(6-S-1) 式就等于

$$(v-b)(F(b^{-1}(b)))^{n-1} \qquad (6-S-2)$$

对于一个竞拍者而言，其应当选择一个 b（记住 b 是一个函数，因而是一个策略）使得 (6-S-2) 式最大化。对 (6-S-2) 式求一阶条件，得

$$-(F(b^{-1}(b)))^{n-1}+\frac{(v-b)(n-1)(F(b^{-1}(b)))^{n-2}F'(b^{-1}(b))}{b'(b^{-1}(b))}=0$$

其中利用了反函数的导数等于原函数的导数的倒数这一法则，即 b^{-1} 的导数等于 $1/b'$。根据 $b^{-1}(b)=v$，又得

$$b'(v)(F(v))^{n-1}+(n-1)b(v)(F(v))^{n-2}F'(v)=(n-1)v(F(v))^{n-2}F'(v)$$

① 由于随机变量为连续随机变量，因而可以不用考虑相等的情况（因为相等时的概率为 0）。

上式的左边实际上等于 $b(v)(F(v))^{n-1}$ 关于 v 的一阶导数，因而对上式两边同时求积分，得

$$b(v)(F(v))^{n-1}=\int_{v_{\min}}^{v}(n-1)x(F(x))^{n-2}F'(x)\mathrm{d}x$$

$$=v(F(v))^{n-1}-\int_{v_{\min}}^{v}(F(x))^{n-1}\mathrm{d}x$$

上式的推导利用到了牛顿—莱布尼茨法则和分部积分法，读者可以在任一本微积分的教科书中查到。

对上述整理后得到

$$b^{*}(v)=v-\frac{\int_{v_{\min}}^{v}(F(x))^{n-1}\mathrm{d}x}{(F(v))^{n-1}} \qquad (6-S-3)$$

这表明如果给定报价函数为公共函数，那么给定（任意类型的）其他人的报价，该（任意类型的）竞拍者根据(6－S－3)式来选择报价是最优反应。由于上面分析的竞拍者可以为任何一人，因而该策略组合就构成了一个对称贝叶斯纳什均衡，其中每一位竞拍者的报价函数为 $b^{*}(v)$。

实际上，如果我们假设 $F(v)$ 为[0,1]上的均匀分布，那么(6－S－3)就等于 $v(1-1/n)$，如果 $n=2$，那么(6－S－3)就等于 $v/2$，与我们前面的分析与推断完全一致。

㊂ 习题 6.16　请证明如果 $F(v)$ 为[0,1]上的均匀分布，那么(6－S－3)等于 $v(1-1/n)$；如果 $n=2$，(6－S－3)等于 $v/2$。

㊥ 习题 6.17　请证明(6－S－3)式中的 $b^{*}(v)$ 是一个增函数。同时证明，当 $v>v_{\min}$ 时，$b^{*}(v)<v$，当 $v=v_{\min}$ 时，$b^{*}(v_{\min})=v_{\min}$，即除了价值最小的竞拍者外，所有其他竞拍者的报价都低于其价值。

一个有趣的问题是：一级价格拍卖的对称贝叶斯纳什均衡的预期收益是否收敛于二级价格拍卖中的优策略组成的贝叶斯纳什均衡？

在二级价格拍卖中，每一个竞拍者的优策略为报价等于价值，并构成一个优策略贝叶斯纳什均衡。如果他中标，那么其实际支付价格为除他之外 $n-1$ 个竞拍者中的最高价值，我们不妨假设其为 v_{second}，显然 v_{second} 是一个随机变量，由于中标者知道 $v_{second}<v$（否则他就不会中标），因而其实际支付价格的期望值就为 $E[v_{second}\mid v_{second}<v]$。$v_{second}$ 的积累分布函数不妨定义为 H，那么 $H(x)=(F(x))^{n-1}$。因而，在小于 v 的条件下，v_{second} 的期望值为

$$E[v_{second}\mid v_{second}<v]=\frac{\int_{v_{\min}}^{v}xH'(x)\mathrm{d}x}{H(v)}$$

$$= \frac{\int_{v_{\min}}^{v} (n-1)x(F(x))^{n-2}F'(x)\,dx}{(F(v))^{n-1}}$$

这实际上就是(6－S－3)式，$b^*(v) = E[v_{second} \mid v_{second} < v]$，即二级价格拍卖的优策略贝叶斯纳什均衡与一级价格拍卖对称贝叶斯纳什均衡带给拍卖者的预期收益完全一样。需要强调的一点是，无论什么拍卖机制，$b < v_{\min}$根本不会出现(你能证明吗?)。因而在一级价格拍卖中，如果竞拍者的价值等于最小值，那么他的报价就只能等于最小价值。实际上，我们始终可以将 v 的最小值设为 0，而不会改变任何结论，就像正态分布的标准化一样。

推论 1 如果每一个竞拍者都为风险中性，其价值(即信号)都独立来自于积累分布函数 $F(v)$，那么贝叶斯密封式二级价格拍卖的优策略均衡与贝叶斯密封式一级价格拍卖的对称均衡的结果完全一样，即拍卖者的预期收益完全一样。

(难) 习题 6.18 现在请你结合上述分析给出等收益定理的一个大致证明。

(二) 公共价值

上面的分析假设竞拍者的价值彼此独立，但实际生活中可能竞拍者的价值不仅依赖于其接收到的信号，而且还依赖于其他竞拍者的信号。我们定义竞拍者 i 的公共价值为 v_i，它是所有信号的增函数。那么，公共价值密封式一级价格拍卖与二级价格拍卖的贝叶斯博弈为 $G_B = \{N, Z, A, T, P, u\}$，其中：

(1) 参与者集合：$N = \{1, 2, \cdots, n\}$，即存在 n 个竞拍者。

(2) 状态集合：$Z = \{z\}$，$z = (t_1, t_2, \cdots, t_n)$，即所有竞拍者的信号向量构成一个状态。

(3) 行动集合：每一个竞拍者 i 的行动就是给出一个报价 $b_i \geq 0$。

(4) 信号集合：对于每一个竞拍者 i 而言，其只能观察到自己的信号，即 $\tau_i(t_1, \cdots, t_i, \cdots, t_n) = t_i$。

(5) 信念：对于每一个竞拍者 i 而言，给定其类型 t_i，其他所有竞拍者的信号都彼此独立。

(6) 偏好：每一个竞拍者 i 的伯努利收益函数为

$$u_i(b_i, (t_1, \cdots, t_n)) = \begin{cases} [v_i(t_1, \cdots, t_n) - p(b_i)]/m & b_i \geq b_j \\ 0 & b_i < b_j \end{cases}$$

其中 b 表示报价，m 表示包括 i 在内有 m 个竞拍者的报价最高。如果 $p(b_i) = b_i$，上述贝叶斯博弈即为一级价格拍卖；如果 $p(b_i) = b_j$，b_j为第二高价，上述博弈则为二级价格拍卖。

1. 二级价格拍卖的贝叶斯纳什均衡

为了简化分析，我们仍然假设 $n=2$，信号 t 服从$[0,1]$上的均匀分布 $F(t)$，但不同的是，每一位竞拍者 i 对标的物的价值判断 $v_i=at_i+ct_j$，$a\geqslant c\geqslant 0$。如果 $c=0$，那么价值就是私人的，如果 $a=c$，那么所有竞拍者的价值都相同，我们称其为**纯公共价值**。例如，北京长安街上的住房如果拿到拍卖所拍卖，那么其价格一定等于市场价，即不会高也不会低，因为竞拍者对房产的价值判断都等于市场价。公共价值分析的核心点在于：虽然竞拍者不能观察到其他竞拍者的信号，但是他可以根据其他竞拍者的报价来获取有用的信息。要直接寻找贝叶斯均衡仍然是困难的，我们不妨猜测每一个竞拍者 i 报价 $b_i=(a+c)t_i$为一个贝叶斯纳什均衡，接下来所需要做的就是证明之。

不失一般性，给定竞拍者 2 的类型为 t_2，并按线性策略报价，那么竞拍者 1 是否也如此呢？为了得到竞拍者 1 的预期收益，我们需要知道其赢得标的物的概率，支付的预期价格及对竞拍者 2 的信号的猜测。

给定竞拍者 2 的报价为 $b_2=(a+c)t_2$，那么只有当 $b_1>(a+c)t_2$时，竞拍者 1 才可能获得标的物。换言之，竞拍者 1 获得标的物的概率为 $Pr(b_1/(a+c)>t_2)$，因为 t_2服从均匀分布，所以 $Pr(b_1/(a+c)>t_2)=b_1/(a+c)$。

其次，如果竞拍者 1 赢得标的物，那么他支付的预期价格为 b_2，可知 b_2 一定小于 b_1（因为竞拍者 1 中标，而 $b_2=b_1$的概率为 0），又因为 t_2为均匀分布，所以 b_2就为服从$[0,b_1]$上的均匀分布，因而竞拍者 2 报价的期望值（基于 $b_2<b_1$的条件）就等于 $b_1/2$，即 $E[b_2\mid b_2<b_1]=b_1/2$。

竞拍者 1 对竞拍者 2 信号的猜测实际上就是信号 t_2的期望值。如果竞拍者 1 赢得标的物，那么什么样的信号 t_2才会给出小于 b_1的报价？给定竞拍者 2 的类型为 t_2，给出的报价$(a+b)t_2$小于 b_1的信号的期望值实际上就等于 $b_1/2(a+c)$。因为 $E[t_2\mid b_2<b_1]$，而 $b_1>b_2\to b_1>(a+c)t_2\to b_1/(a+c)>t_2$，又 t_2为$[0,1]$上的均匀分布，所以其期望值 $E[t_2\mid b_2<b_1]=b_1/2(a+c)$。

综上所述，竞拍者 1 如果赢得标的物，那么其预期收益函数为

$$(v_1-b_2)Pr(b_1>b_2)+0\times Pr(b_1<b_2)$$

即
$$\frac{a}{2(a+c)^2}[2(a+c)t_1-b_1]b_1$$

竞拍者 1 所要做的就是选择 b_1使得上式最大化。容易证明，当 $b_1=(a+c)t_1$时，竞拍者 1 的预期收益最大。利用对称性，可知线性策略为公共价值二服价格拍卖的一个贝叶斯均衡。这时，每一个竞拍者的价值不仅是其信号的函数，也是其他竞拍者信号的函数。

㊊ 习题 6.19　请你考虑 $n>2$ 的情景，线性策略是否仍是二级价格拍卖的贝叶斯纳什均衡？

需要注意的是，竞拍者如果仅仅根据自己的信号及对其他竞拍者的信号的猜测来决定其价值，那么有可能会作出过高的价值判断，从而给出过高的报价，这被称为**赢者的诅咒**（winner's curse）。因而，现实中，一个竞拍者在形成自己的价值时，有可能并不仅仅基于所有竞拍者的信号，而是会考虑其他的信息。例如，在股市上，所谓的追涨杀跌，就是一种赢者的诅咒（赢者是指买到股票的人），而聪明的投资者不仅仅会看其他投资者如何做，还会考虑其他的重要信息，例如公司的基本面、行业前景、国家的宏观环境等因素。

2. 一级价格拍卖的贝叶斯纳什均衡

在私人价值中，我们曾证明每一个竞拍者的报价为 $v/2$ 构成一个贝叶斯均衡，那么在公共价值中，如果每一个竞拍者的报价为 $(a+c)t/2$ 是否为一个贝叶斯纳什均衡？所有的假设仍与前面一样。

不失一般性，给定竞拍者 2 的报价为 $(a+c)t_2/2$，那么什么是竞拍者 1 的最优反应？要得到竞拍者 1 的预期收益，我们需要求出竞拍者 1 赢得标的物的概率，以及对竞拍者 2 的信号的猜测。

给定竞拍者 2 的报价为 $b_2=(a+c)t_2/2$，那么只有当 $b_1>(a+c)t_2/2$ 时，竞拍者 1 才可能获得标的物。这就意味着，竞拍者 1 获得标的物的概率为 $Pr(2b_1/(a+c)>t_2)$，因为 t_2 服从均匀分布，所以 $Pr(2b_1/(a+c)>t_2)=2b_1/(a+c)$。

其次，如果竞拍者 1 赢得标的物，那么什么样的信号 t_2 才会使竞拍者 2 给出小于 b_1 的报价？如果竞拍者 2 的类型为 t_2，那么其期望值等于 $E[t_2 \mid b_2<b_1]$，因为 $b_1>b_2 \to b_1>(a+c)t_2/2 \to 2b_1/(a+c)>t_2$，又因为 t_2 为 $[0,1]$ 上的均匀分布，所以竞拍者 2 给出报价小于 b_1 的信号的期望值实际上就等于 $b_1/(a+c)$，即 $E[t_2 \mid b_2<b_1]=b_1/(a+c)$。因而，竞拍者 1 的预期收益为

$$(v_1-b_1)Pr(b_1>b_2)+0\times Pr(b_1<b_2)$$

即等于

$$\frac{2a}{(a+c)^2}[(a+c)t_1-b_1]b_1$$

竞拍者 1 所要做的就是选择 b_1 以使得上式最大化。容易证明，当 $b_1=(a+c)t_2/2$ 时，竞拍者的预期收益最大，因而 1/2 报价就是一个最优策略。利用对称性，可知 1/2 报价为一个贝叶斯纳什均衡。下面的习题要求读者给出 $n>2$ 时的证明。

㊥ *习题 6.20　请你考虑 $n>2$ 的情景，$1-1/n$ 报价是否是公共价值一级价格拍卖的贝叶斯纳什均衡？如果 $F(t)$ 为任意分布呢？*

实际上，即使在公共价值的情况下，二级价格拍卖与一级价格拍卖的预期收益同样是相等的，即一级价格拍卖和二级价格拍卖中的竞拍者的报价都为公共

报价(即策略都一样),那么两种拍卖机制下拍卖者在对称均衡下的预期收益(在极限上)都相等。这里我们将不加证明地给出公共价值下的等收益定理。

命题 6.2 公共价值等收益定理(Myerson, 1981)给定每一位竞拍者① 风险中性;② 信号彼此独立;③ 公共价值,即竞拍者的价值 v 不仅依赖自己接受到的信号,而且还取决于其他竞拍者的信号。对于任何拍卖机制,其中的对称贝叶斯纳什均衡满足:① 拍卖品为最高信号者得,② 最小可能价值的竞拍者的期望支付为 0。那么在这些拍卖机制的对称均衡中,拍卖者所得的预期收益完全一样。

不过,在更一般的情况下,各种拍卖机制还是存在差异的,米尔格罗姆(Milgrom)和韦伯(Weber)在 1982 年证明了这个结果[①]。

命题 6.3 拍卖机制有下列关系:

(1) 荷式拍卖和密封式一级价格拍卖完全等价;

(2) 在英式拍卖中,拍卖者的预期收益大于密封式二级价格拍卖;

(3) 在密封式二级价格拍卖中,卖方的预期收益大于密封式一级价格拍卖。

由命题 6.3 可知,英式拍卖是四种拍卖机制中,预期收益最高的一种拍卖机制。这就是为什么在一般拍卖所里,最常见到英式拍卖。既然如此,所有拍卖都采用英式拍卖岂不更好? 实际上,拍卖制度的根本目的在于发现拍卖品的“真实价值”,这就是为什么我们不会把轿车、电视、大白菜拿去拍卖。其次,防止竞拍者之间的勾结也非常重要。在拍卖的过程中,常会发生竞拍者相互勾结的情形。这以政府经常性的招标最常见。例如,在工程招标中,营造公司相互串谋决定投标价格,然后以非常高的造价让某一公司中标。但由于投标勾结者之间的约定没有法律效力,也没有很强的约束力,因此这种勾结要成功必须借助拍卖制度本身的缺陷。如果投标过程使用英式拍卖,那么任何一家厂商想要违反约定将非常困难,因为其他厂商能当场发现该厂商违反事前约定,而用更低的价格来反制,使得违反约定的图谋难以得逞。正因为如此,英式拍卖反而增强了勾结的可能性。但在密封制度下,由于厂商无法知道其他厂商的标价,因此勾结较难实现。从竞拍者本身来看,勾结本身也需要一些条件。较有可能产生勾结的拍卖市场是竞标者数目较少且经常举行拍卖活动的场所,因为达成和维持一个勾结协定的难度随竞标者数目的增加而成几何数增加,这就是所谓的十个监督者不如一个竞争者。在工程招标中,一则由于营造厂商的数目较少,二则由于工程招标活动较为频繁,因此勾结的可能性很大。这就是公共工程招标多采用密封式

① Milgrom P. and Weber R. J. “A Theory of Auctions and Competitive Bidding”. *Economitrica*. 1982, 50: 1089 – 1122

拍卖制度的原因。

五、有效率不一定有交易

下面我们分析一个叫做双向拍卖的交易博弈。买方给出一个买价 p_b，卖方同时给出一个卖价 p_s。如果 $p_b \geq p_s$，则交易以 $p=(p_b+p_s)/2$ 的价格进行，如果 $p_b<p_s$，则不发生交易（股票市场就是这样交易的）。买方对标的物的估价为 v_b，卖方的估价为 v_s，双方的估价都是私人信息，并且服从区间[0,1]上的均匀分布。如果买方以价格 p 购得商品，则可获得 $v_b—p$ 的收益，卖方获得 $p—v_s$ 的收益；如果交易未成功，买卖双方收益为0。

㊣ 习题6.21　请你给出双向拍卖交易博弈的基本式。

根据行动空间和类型空间定义买卖双方的策略，买方的策略定义为函数 $p_b(v_b)$，即在每一个可能的类型下将会给出的买价。类似地，卖方的策略为 $p_s(v_s)$。

不失一般性，给定卖方的策略为 $p_s(v_s)$，那么如果交易发生，其概率一定为 $Pr(p_b(v_b)\geq p_s(v_s))$。其次，在 $p_b(v_b)\geq p_s(v_s)$ 的条件下，买者需要对卖者的出价进行猜测，它实际上等于 $E[p_s(v_s)\mid p_s(v_s)\leq p_b]$，即在卖方价格小于买方价格的条件下，买方对卖方价格的期望值。因而买者的预期收益为

$$\max_{p_b}\left[v_b-\frac{p_b+E[p_s(v_s)]\mid p_b\geq p_s(v_s)}{2}\right]Pr\{p_b\geq p_s(v_s)\}$$

(6－E－1)

任意类型为 v_b 的买者所要做的就是选择 p_b 使得上式最大化。

同理，给定买方的策略为 $p_b(v_b)$，对于任意类型为 v_s 的卖者其预期收益为

$$\max_{p_s}\left[\frac{p_s+E[p_b(v_b)\mid p_b(v_b)\geq p_s]}{2}-v_s\right]Pr\{p_b(v_b)\geq p_s\}$$

(6－E－2)

任意类型为 v_s 的卖者所要做的就是选择 p_s 使得上式最大化。

如果 $[p_b^*(v_b),p_s^*(v_s)]$ 为式(6－E－1)和(6－E－2)的解，那么该策略组合就是交易博弈的贝叶斯纳什均衡。

由于出价函数 p 可能有无穷多个，因而交易博弈的贝叶斯均衡可能有无穷多个。这里我们考虑两个特殊的均衡：单一价格均衡和线性均衡。所谓单一价格均衡，就是要么以单一的价格成交，要么不成交。设 x 是属于区间[0,1]上的任意一个值，令买方的策略为：如果 $v_b\geq x$，则出价 x，其他情况下出价0；同时令卖方的策略为：如果 $v_s\leq x$，则卖价为 x，其他情况下卖价为1。

给定买方的策略，容易证明，在 $v_s \leq x$ 下成交，在 $v_s > x$ 下不成交，是卖方的严格优策略，所以在 $v_s \leq x$ 时卖价为 x，其他情况下卖价为 1，是卖方对买方策略的最优反应。反之，给定卖方的策略，在 $v_b \geq x$ 下成交，在 $v_b < x$ 下不成交，是买方的严格优策略，所以在 $v_b \geq x$ 时出买价 x，其他情况下出价 0，是买方对卖方策略的最优反应。因而，买卖双方都采用单一价格策略为该博弈的一个贝叶斯均衡。图 6－7 给出了在这一均衡结果下所有可能发生交易的区域，为阴影部分。

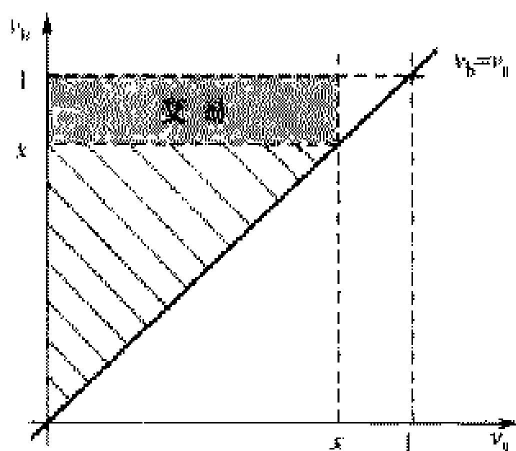


图 6－7　单一价格均衡

从微观经济学的知识来讲，如果 $v_b > v_s$，那么交易都应当发生，因为这时交易能够改善双方的福利（例如以 $(v_b + v_s)/2$ 的价格成交），因而交易是有效率的。但图中斜线部分虽满足效率条件，却没有发生交易。

现在我们推导交易博弈的线性均衡。由于估价服从均匀分布，线性均衡是存在的。除了单一价格均衡和线性均衡之外，博弈还存在着许多其他的均衡，但线性均衡有着良好的效率特征。

假设卖方的策略为 $p_s(v_s) = a_s + c_s v_s$，则 p_s 服从区间 $[a_s, a_s + c_s]$ 上的均匀分布，于是（6－E－1）式可化为

$$\max_{p_b}\left[v_b - \frac{1}{2}\left(p_b + \frac{a_s + p_b}{2}\right)\right]\frac{p_b - a_s}{c_s}$$

由于 p_s 服从均匀分布，在 $p_b \geq p_s(v_s)$ 的条件下，p_s 的最小值为 a_s，最大值为 p_b，所以 $E[p_s(v_s) \mid p_b \geq p_s(v_s)] = (a_s + p_b)/2$。

对上式求一阶条件可得

$$p_b = \frac{2}{3}v_b + \frac{1}{3}a_s \qquad (6-E-3)$$

相似地，假设买方的策略为 $p_b(v_b) = a_b + c_b v_b$，则 p_b 服从区间 $[a_b, a_b + c_b]$ 上的均匀分布，于是（6－E－2）式可化为

$$\max_{p_s}\left[\frac{1}{2}\left(p_s+\frac{p_s+a_b+c_b}{2}\right)-v_s\right]\frac{a_b+c_b-p_s}{c_b}$$

由于 p_b 服从均匀分布，在 $p_b(v_b)\geq p_s$ 的条件下，p_b 的最小值为 p_s，最大值为 a_b+c_b，所以 $E[p_b(v_b)\mid p_b(v_b)\geq p_s]=(p_s+a_b+c_b)/2$。

求上式的一阶条件可得

$$p_s=\frac{2}{3}v_s+\frac{1}{3}(a_b+c_b) \tag{6-E-4}$$

这表明，如果卖方选择一个线性策略，则买方最优的策略也是线性的，反之亦然，即存在一个线性均衡。

由(6－E－3)式可知 $a_b+c_bv_b=p_b=\frac{2}{3}v_b+\frac{1}{3}a_s$，从而得 $c_b=2/3$，$a_b=a_s/3$。

由(6－E－4)式可知 $a_s+c_sv_s=p_s=\frac{2}{3}v_s+\frac{1}{3}(a_b+c_b)$，从而得 $c_s=2/3$，$a_s=(a_b+c_b)/3$。

因而具体的线性均衡策略为

$$p_b(v_b)=\frac{2}{3}v_b+\frac{1}{12} \tag{6-E-5}$$

和

$$p_s(v_s)=\frac{2}{3}v_s+\frac{1}{4} \tag{6-E-6}$$

[(6－E－5)，(6－E－6)]即为交易博弈的一个贝叶斯纳什均衡。

双向拍卖中交易发生的必要条件是 $p_b\geq p_s$。如图 6－8 所示，当卖方的估价高于 3/4 时，他的要价超过了买方的最高可能出价 $p_b(1)=3/4$；当买方的估价低于 1/4 时，他的出价低于卖方的最低可能要价 $p_s(0)=1/4$。

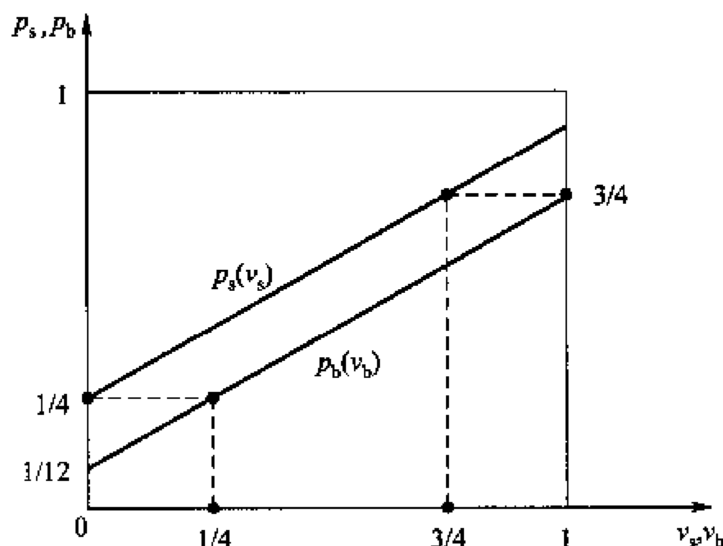


图 6－8　线性均衡下的可能交易

在上述两种情况下，交易都不会发生。这意味着，在线性均衡中，只有当 $v_b \geqslant v_s + (1/4)$ 时才会有交易发生，如图 6－9 所示。

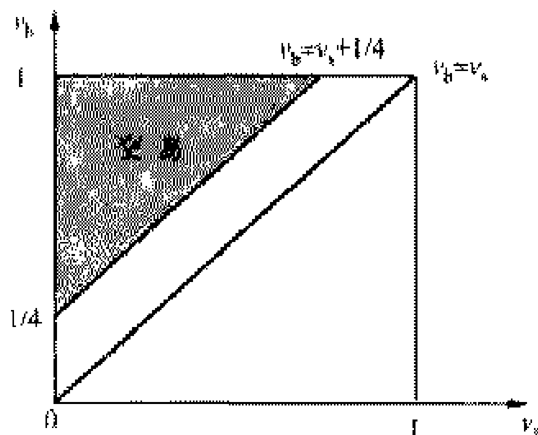


图 6－9　线性均衡下的交易区间

比较图 6－7 和图 6－9，可清楚地看出在单一价格均衡及线性均衡中，交易发生所要求的估价组合。在这两种情况下，交易的潜在价值最大时（具体讲，当 $v_s=0$，$v_b=1$），都会发生交易。但是，单一价格均衡漏掉了一些有价值的交易，比如 $v_s=0$，$v_b=x-\varepsilon$，其中 ε 是足够小的正数，而包含了一些价值不大的交易，比如 $v_s=x-\varepsilon$ 且 $v_b=x+\varepsilon$。相反，在线性均衡中，过滤掉了所有价值不大的交易，只包含了价值至少在 1/4 以上的交易。这表明从参与者可得到的期望收益的角度来看，线性均衡要优于单一价格均衡，但是否是最优还需要考察其他的均衡，幸运的是线性均衡确实是最优的。

在类型服从均匀分布的情况下，线性均衡不仅存在，而且相对于其他均衡是最优的。迈尔森（Myerson）和萨特思韦特（Satterthwaite）在 1983 年证明了这一结论：对于这里考虑的估价的均匀分布，双向拍卖的线性均衡中参与者的期望收益要高于其他任何贝叶斯均衡。这意味着在双向拍卖中，不存在这样的贝叶斯均衡：交易当且仅当有效率时（$v_b \geqslant v_s$）必定发生（因为线性均衡过滤掉了一些价值不大的交易）。他们进一步证明了这一结果的普遍性，即无效率定理。

定理 6.2　无效率定理如果 v_b 在区间 $[x_b, y_b]$ 连续分布，v_s 在区间 $[x_s, y_s]$ 连续分布，其中 $y_s > x_b$ 且 $y_b > x_s$，那么在讨价还价博弈中，仅当有效率时（$v_b \geqslant v_s$），贝叶斯纳什均衡的交易并不必然产生，有效率只是贝叶斯均衡交易的必要条件。

无效率定理修正了微观经济学中有关一般均衡的命题。在微观经济学中，只要交易是有效率的，我们就能得到一条契约线，从而保证一般均衡出现。但根据无效率定理，如果经济中存在着私人信息（信息不完全），那么交易有效率并不能保证交易一定发生。产生有效率而无交易的根本原因就在于效率和激励在绝大多数情况下就像熊掌与鱼肉不可兼得。

第四节　显示原理

在前面讲述拍卖理论的时候，我们已经看到：① 拍卖机制有多种。② 明确了使用某种拍卖机制后，拍卖中的投标者的类型服从什么样的概率分布，策略空间是不是非常庞大。这些问题十分关键。③ 即使以上都明确，也还存在其他的一些可能性，例如，投标者可能还需要支付一定的进入费用，标的物可能有底价，标的物可能卖不出去，卖出去也不总是卖给出价最高的人，等等。这意味着，要找到最优的均衡解，我们必须：首先，依次考虑所有可能的机制，确定每一个机制可能出现的均衡解；其次，比较所有不同机制下的均衡解，找出全局性的最优均衡；最后，要考虑可能出现的一些情况是否会对均衡解造成扭曲。以上这些因素表明，要找到贝叶斯博弈中的最优均衡解与大海捞针无异。那么有没有什么办法来最大程度地简化我们所面临的贝叶斯博弈呢？

一个天真的想法是，如果人人都是讲真话的君子那该多好！只要人们把他们的真实类型报告出来，所有的问题都迎刃而解。但在理性人的假设下，一个自然的诘问是，如果报告真实偏好不能带来好处，为什么要真实报告自己的偏好？如果竞标者真实报告自己的偏好就意味着要付更多的钱，学生真实报告自己的不努力就意味着要求老师网开一面没了理由，工人真实报告自己工作偷懒就意味着加薪泡汤，企业主真实报告收入就意味着税收翻番发财成空想。这就给我们一个提示：人们不会主动地报告自己的真实偏好，而必须设计一个机制（或制度），在这个机制中参与者不得不报告他的真实偏好。那么能不能设计这样的机制呢？迈尔森（Myerson）1979 年提出的显示原理证明这样的机制是存在的[①]。显示原理不仅仅运用于拍卖机制（拍卖机制只是一个例子），而且可以运用在所有的私人信息博弈，例如公共产品配置、外在性解决、垄断监管、道德风险、逆向选择等问题中。这表明显示原理具有很好的理论意义。这里仍然以拍卖机制为例来阐述显示原理。

考虑拍卖者希望设计一个拍卖以使他的期望收益最大化，显然逐一列出卖方可能考虑的不同的拍卖机制，并求最优贝氏均衡是不可能的（例如类型服从不同的概率分布就可能使策略空间无穷大）。这时拍卖者就可以借助显示原理，从两个方面来简化他的问题。

第一，拍卖方可以将其分析集中于下面的博弈：

（1）投标者同时报告（也许是不诚实的）他们自己的类型。投标者 i 的报告

① Myerson R. "Incentive Compatibility and the Bargaining Problem". *Econometrica*. 1979, 47: 61 ~ 73

函数定义为 $r_i:T_i\to T_i$，它是从类型空间到类型空间的函数，记作 $r_i(t_i)$，其中 t_i为参与者 i 的真实类型。如果参与者 i 讲真话，那么 $r_i(t_i)=t_i$。

（2）给定投标者的报告组合$(r_1,r_2,\cdots,r_n)$，投标者 i 报价 $x_i(r_1,r_2,\cdots,r_n)$ 时，将以 $q_i(r_1,r_2,\cdots,r_n)$ 的概率获得标的物。对于所有可能的报告组合$(r_1, r_2,\cdots,r_n)$，概率$\sum_{i=1}^{n}q_i(r_1,\cdots,r_n)\leqslant 1$。

我们把参与者报告自己偏好的贝叶斯博弈称为**直接机制**。

第二，根据显示原理，存在策略组合$(r_1(t_1)=t_1,r_2(t_2)=t_2,\cdots,r_n(t_n)=t_n)$是直接机制中的一个贝叶斯纳什均衡。这表明，通过设计适当的支付函数和概率函数$\{x_1(r_1,\cdots,r_n),\cdots,x_n(r_1,\cdots,r_n);q_1(r_1,\cdots,r_n),\cdots,q_n(r_1,\cdots,r_n)\}$，可以使每一个参与者的最优策略是实话实说，即对 $\forall t_i\in T_i, r_i(t_i)=t_i$。实话实说成为贝叶斯纳什均衡策略表明实话实说是参与者的最优策略，每一个参与者都不会偏离它（说假话）。

上述叙述的关键是，实话实说要成为直接机制中的贝叶斯纳什均衡策略必须满足激励兼容条件。**激励兼容**条件通俗地说，就是委托人（拍卖者）希望代理人（竞标者）采取的策略（行动）使得委托人的利益最大，同时也使得代理人本身的利益最大。例如，一位画家（委托人）拍卖他的作品，他当然希望拍卖的价格越高越好，如果他委托一位拍卖师（代理人）帮他拍卖，并规定拍卖师的佣金按拍卖价格的百分比支付，那么拍卖师的策略就满足激励兼容条件，因为拍卖师拍卖的价格越高他得到的佣金越多，画家得到的收益也越多。一个企业主和工人同样形成了一种委托代理关系，如果企业主付给工人的工资是一个固定数，企业主很难监督工人劳动（存在私人信息），那么工人就可能偷懒，使得企业主受损、工人获益（偷懒使效用更大），这时工人的策略（行为）就不满足激励兼容条件。

定理 6.3　显示原理任何一个贝叶斯博弈都可重新表述为一个满足激励兼容的直接机制，它的任何贝叶斯纳什均衡等价于该直接机制的贝叶斯纳什均衡。

由于任何贝叶斯博弈都对应着一个直接机制，因而我们可以用显示原理来简化对贝叶斯博弈的分析，只要分析直接机制就可以了。对于直接机制不存在的特性，不能实现的均衡，相应的贝叶斯博弈也必然不存在这种特性、不能实现这种均衡。例如，根据显示原理我们知道要实现一个对拍卖者最理想的拍卖机制（等同于一个直接机制），它必须满足激励兼容的条件，光有效率是远远不够的，这就（不太严格）证明了我们前面提到的无效率定理——有效率不一定有交易，有效率只是贝氏均衡交易的必要条件。

为了给出显示原理的正式表述及其证明，我们假设存在一个贝叶斯静态博

弈 $G_B=\{I,A,T,p,u\}$，其中 $s^*=(s_1^*,\cdots,s_n^*)$ 为它的一个贝叶斯纳什均衡。我们将构建一个有实话实说均衡的直接机制来重新表示 s^*。符合条件的直接机制是一个贝叶斯静态博弈，和 G_B 有相同的类型空间和推断，但其行动空间和收益函数却不一样。由于直接机制中参与者 i 的行动就是（可能是不诚实地）报告他的所属类型，因而直接机制中参与者 i 的行动空间等于类型空间 T_i。新的收益函数则要复杂一些，它们不仅决定于原博弈 G_B，还决定于原博弈的均衡 s^*。问题的关键在于运用 G_B 的均衡解 s^* 来保证实话实说是直接机制的均衡解，具体证明步骤如下。

s^* 是 G_B 的一个贝叶斯纳什均衡，意味着对每一参与者 i，s_i^* 为 i 对其他参与者策略 $(s_1^*,\cdots,s_{i-1}^*,s_{i+1}^*,\cdots,s_n^*)$ 的最优反应。更为具体地，对 i 的每一种类型 $t_i\in T_i$，$s_i^*(t_i)$ 都是 i 对给定其他参与者的策略为 $(s_1^*,\cdots,s_{i-1}^*,s_{i+1}^*,\cdots,s_n^*)$ 时，从 A_i 中选择的最优行动。那么，如果 i 的类型为 t_i，并且我们允许 i 从 A_i 中包含 $s_i^*(t_i)$ 的一个子集中选择行动（记住是子集），仍然假定其他参与者的策略为 $(s_1^*,\cdots,s_{i-1}^*,s_{i+1}^*,\cdots,s_n^*)$，则 i 的最优选择仍为 $s_i^*(t_i)$。选定直接机制中的收益函数时，就是设法使每一参与者面临和上面相同的选择。

确定直接机制中收益函数的方法如下：首先将新博弈中参与者关于类型的报告 $r=(r_1,\cdots,r_n)$ 代入原博弈中的均衡策略 s^*，然后将求得的原博弈的均衡行动 $s^*(r)=(s_1^*(r_1),\cdots,s_n^*(r_n))$ 代入原博弈的收益函数，从而得到直接机制的收益函数，用公式表示，i 的收益函数为

$$v_i^*(r,t)=u_i[s^*(r),t]$$

其中 $t=(t_1,\cdots,t_n)$。对于这样的收益函数，我们可以理解为存在一个计算机系统，这个计算机系统是按直接机制设计的，它进行如下的运算：

(1) 所有参与者同时报告自己的类型，并且将它输入计算机中；

(2) 计算机根据每一个参与者的报告 r_i，代入原博弈参与者的均衡策略 s_i^*，计算出如果参与者的类型为 r_i 时，他在均衡 s^* 中将会如何行动——也就是 $s_i^*(r_i)$。最后，计算机根据博弈规则计算出所有参与者在行动 s^* 下各自的收益，并将收益自动记入每一个参与者的账户中。

显然在这个由计算机系统组成的直接机制中，所有参与者唯一的行动就是报告自己的类型（将自己的类型输入计算机），然后参与者得到各自的收益，博弈随之结束。毫无疑问，这个系统是一个典型的贝叶斯静态博弈。它可以表示为 $G_D=\{N,R(T),P,v\}$，其中 $R_i(T)$ 为报告空间，它实际上等于类型空间 T_i，v_i 为直接机制中的收益函数，它依赖于参与者的行动（报告）和真实类型。

最后，我们要证明 $r(t)=t$ 是上述直接机制中的贝叶斯纳什均衡。通过向计算机报告自己的类型 $r_i(t_i)\in T_i$，参与者 i 事实上等于从 A_i 中选择了行动 $s_i^*(r_i)$。如果所有其他参与者都真实报告自己的类型 $r_{-i}(t_{-i})=t_{-i}$，则等于选择了策略 $s_{-i}^*=(s_1^*,\cdots,s_{i-1}^*,s_{i+1}^*,\cdots,s_n^*)$。而前面我们已经知道，$s^*$ 是原博弈的贝叶斯纳什均衡，如果其他参与者选择了这样的策略，当 i 的类型为 t_i 时，参与者 i 的最优行动只能是 $s_i^*(t_i)$。这意味着，如果其他参与者真实报告自己的类型，则参与者 i 的最优策略也只能是真实报告自己的类型。现在设 i 为参与者集合 N 中的任意一个值，即 $\forall i\in I$，上述结论同样成立，即实话实说是直接机制中的一个贝叶斯纳什均衡。更为正式的，在静态贝叶斯博弈 $G_D=\{N,R(T),P,v\}$ 中，对 $\forall t_i\in T_i$，每一参与者 i 选择实话实说策略 $r_i(t_i)=t_i$ 构成一个贝叶斯纳什均衡。

显示原理最重要的理论意义就在于它表明在广泛的社会选择中存在着不可能性定理，著名的阿罗不可能性定理只不过是不可能性定理在社会公共选择下的特例。所谓不可能性定理是指，在每一个人都平等自由的条件下，各自根据自身的偏好理性地作出选择，然后通过我们所能想到的任何民主制度来合成个人选择以形成统一的社会选择，那么这个社会选择必然是无效率的，也就是效率和民主制度是冲突的。正像显示原理所揭示的任何存在私人信息的制度（直接机制和间接机制）都可表示成说真话做实事的直接机制，然而讲真话做实事并不能给经济人带来利益的最大化，要让人们讲真话做实事就必须给予激励，而要实现这种激励就必然以牺牲效率为代价。

在什么情况下人类能摆脱不可能定理的羁绊？只有在类似于完全竞争这样的制度结构中不可能定理才会失效。完全竞争的一个重要条件是市场参与者的人数趋于无穷大，因而从这个意义上来讲，普选＋直选是民主制度的最高形式。美国2000年总统大选出现的由法官来决定总统的争端实际上只不过是不可能性定理在社会政治选择中时不时体现自己存在的一种表现而已。那么为什么没有一个国家的民主制度能够达到普选＋直选？根本原因在于普选＋直选存在着巨大的交易费用，而交易费用的产生与信息不完全存在着极大的联系，而在完全竞争下信息是完全的，交易费用为零，而这在现实中不过是幻想而已。这使我们不得不想起在公共地悲剧中哲学家怀海特所说的话：问题的解决不在于技术而在于精神。人类物质技术的不断进步并不是人类文明进步的充分条件，人类各种难解之症的解决与其说依赖于技术，不如说依赖于人类崇高无上的精神和思想。

关键词

豪尔绍尼转换　状态　状态类 t_i　信号函数　信念　贝叶斯博弈的基本式　静态贝叶斯博弈中的纯策略(混合策略)　贝叶斯(混合)纳什均衡　私人价值　公共价值　英式拍卖　荷式拍卖　(一级和二级)密封式拍卖　等收益定理　无效率定理　直接机制　激励兼容　显示原理

内容要点

1. 贝叶斯博弈中的一个困难就是存在不同的状态,不可能直接分析。解决这个问题的一个方法就是规定不同的状态服从一定的概率分布,这种方法被称为豪尔绍尼转换。通过豪尔绍尼转换,贝叶斯博弈的收益变成了期望收益,通过比较不同策略组合下的期望收益,从而找出贝叶斯博弈中的纳什均衡。

2. 贝叶斯博弈的基本式与完全信息静态博弈相比多了三个要素:即状态集合 Z、信号集合(类型集合)T 和信念 P,因而贝叶斯博弈可表述为 $G_B=\{N,A,Z,T,P,u\}$,策略式博弈实际上是贝叶斯博弈的特例。在贝叶斯博弈中,信息不完全可表示为不同的状态发出的信号一样,既所谓的状态类。同时,参与者对不同状态的信念是根据贝叶斯法则计算出来的,也正因为如此非完全信息博弈被称为贝叶斯博弈。

3. 实际上,通过适当地定义参与者、行动空间、收益函数,贝叶斯纳什均衡完全等价于一个新的策略式博弈的纳什均衡,即命题 6.1。命题 6.1 还充分说明了处理贝叶斯博弈的基本思路(也是唯一的方法),就是通过豪尔绍尼转换将非完全信息博弈变成完全信息博弈来处理。

4. 贝叶斯博弈的一个重要运用就是对拍卖机制的研究,而且拍卖机制也很好地概括了贝叶斯博弈的主要内容。拍卖机制中的等收益定理具有很广的实用性,不论是在私人价值还是公共价值情况下,等收益定理都是成立的。通过对拍卖制度的研究得出的无效率定理修正了传统微观经济学中的一般均衡理论,有效率并不一定有交易,有效率只是交易发生的必要条件而不是充分条件。

5. 根据显示原理可以简化对贝叶斯博弈的分析,同时也说明了激励兼容与效率两者就像鱼与熊掌不可兼得。实际上,如何最大程度地平衡激励兼容与效率的关系构成了信息经济学和机制设计理论的核心,也是各种大大小小的社会制度安排最需要克服的首要问题。

进一步阅读

最早给出贝叶斯博弈概念和思想的是豪尔绍尼(Harsanyi,1967,1968);对贝叶斯博弈进行纯技术性介绍的可参阅 Osborne and Rubinstein(1994);对于显示原理感兴趣的则可以参阅 Myerson(1985),其对显示原理作了较为详细和通俗的介绍。另外,Osborne(2004,Chapter 9)在贝叶斯博弈的框架下,对公共知识的不同结构(例如参与者1不知道参与者2是否知道参与者1的收益函数)对均衡产生的影响作了非常通俗的说明和分析。

第七章

非完全信息动态博弈

在这一章中，我们将会接触到博弈论中最为复杂的模型——非完全信息动态博弈，也称动态贝叶斯博弈。序列均衡是非完全信息动态博弈中的核心概念，由克瑞普斯(Kreps)和威尔逊(Wilson)在1982年最早提出，并发展成为博弈论中最一般的均衡概念。序列均衡是对完美贝叶斯均衡的再精炼，而完美贝叶斯均衡则可以理解为是贝叶斯纳什均衡和子博弈完美均衡的综合。此外，序列均衡也是信息经济学的分析基础。

在这一章中，我们首先给出一般化的扩展式，所谓一般化是指它可以描述各种各样的博弈，无论是静态还是动态，无论是信息完全还是不完全。第二节则围绕着序列均衡展开，序列均衡最突出的特征就是它揭示出，在动态贝叶斯博弈中，所谓的"策略"实际上是"万全行动计划 + 信念系统"，在博弈论中被称为审时度势(assessment)。第三节则是经典举例。序列均衡虽然具有较广的适用性，但在某些情况下仍不够严格，会包含着一些不合理的均衡，因而第四节主要讨论对序列均衡的再精炼。最后一节则是对所有均衡概念作一小结。

客观而言，直到今天理论界仍没有找到能适用于所有非完全信息动态博弈的均衡概念，也缺乏一个普遍适用的方法来解动态博弈的序列均衡。通常，求解序列均衡，除了技术之外，往往还需要一点灵感加运气。

第一节 一般扩展式

在完全信息动态博弈中，要描述一个博弈需要四个要素，即参与者集合、全历史集合、参与者函数和偏好。如果其中加入了不确定性(即信息不完全)，那么仅有这四个要素显然无法准确描述出博弈的特征，而需要一些新的要素。

(1) 自然(nature)。在动态贝叶斯博弈中，有一位参与者参与博弈，但却没有自己的收益，其作用就是决定不同历史发生的概率，我们把这样的参与者称为

自然。实际上,贝叶斯博弈中的豪尔绍尼转换就是由自然来完成的。同时,自然行动的概率分布用概率函数 $\rho: I_0 \times A_0 \rightarrow [0,1]$ 来表示,其中 0 表示自然,I_0 则表示自然的信息集,A_0 表示自然的行动空间。

(2) 信息集(information set)和信息分割(partitioning)。所谓参与者 i 的一个信息集就是指真子历史的一个集合,它具有如下几个特征:① 信息集中的每一个真子历史之后都轮到参与者 i 行动。如果用数学表达,设该信息集为 I_{ik}(即参与者 i 的第 k 个信息集),那么 $\forall h \in I_{ik}$(即 h 为 I_{ik} 中的任一真子历史),必有 $P(h)=i$,$P(\cdot)$ 为参与者函数。② 如果信息集不是只包含一个真子历史,而是包含多个,那么参与者 i 分不清其到底位于哪一个真子历史。这就意味着参与者 i 的行动函数应该完全一样,即对于任意两个历史 $h \in I_{ik}$ 和 $h' \in I_{ik}$,$A_i(h) = A_i(h')$。如果不一样,那么参与者 i 就能区分出其到底位于哪一个真子历史,这就与信息集的要求相违背。这也说明,我们可以将行动函数定义在信息集上,即 $A(I_{ik}) = A_i(h)$。③ 如果把轮到参与者 i 行动之前的所有真子历史看作一个集合 SH_i,即 $\forall h \in SH_i$,$P(h)=i$,那么参与者 i 的任一信息集 I_{ik} 实际上就是 SH_i 的一个子集。如果把参与者 i 的所有信息集集合定义为 I_i,即 $I_i = \{I_{i1}, I_{i2}, \cdots, I_{iK}\}$,称为参与者 i 的信息空间,那么对于任一参与者 i 的信息空间 I_i,我们要求它必须满足两个条件:一是对于 I_i 中的任意两个信息集,两两相交必须为空,二是所有信息集之并等于 SH_i,这在数学上被称为一个分割(partitioning)①。实际上,动态贝叶斯博弈中参与者的不同信息状况就可表示为不同的信息分割。为了方便,以后我们就用**信息分割**来强调信息集的这种特性,即一种信息分割就代表着一种信息状态。例如,在动态博弈中,如果每一个参与者的每一个信息集都只包含一个真子历史,那么这个博弈就是一个完美信息动态博弈。如果用信息分割来表示,就是 $SH_i = \{h_1, \cdots, h_k\}$,$I_{i1} = \{h_1\}, \cdots, I_{ik} = \{h_k\}$,对每一个参与者 i 成立,k 为任意自然数。

(3) 定义在博彩(lottery)上的偏好。经济学中的博彩就是指一个概率分布下的收益。例如,小刘购买了海尔的股票 10 份,那么他有 1/3 的可能获得 100% 的收益,但也有 2/3 的可能获得 0% 的收益,这是一个博彩,不妨定义为博彩 1;如果小刘把钱拿去买体育彩票,那么他有 1% 的可能获得 1 000% 的收益,但有 99% 的可能获得 -100% 的收益(即输掉本金),这是另一个博彩,不妨定义为博彩 2。那么小刘是偏好博彩 1,还是博彩 2 呢?显然,一个合理的想法就是

① 所谓分割就是设 Y 为一个集合,$Y_1, Y_2, \cdots, Y_n$ 为集合 Y 的子集,如果 $Y_1 \cap Y_2$,$Y_2 \cap Y_3$ 等任意两两相交为空集,$Y_1 \cup Y_2 \cup \cdots \cup Y_n = Y$,那么就称 $Y_1, Y_2, \cdots, Y_n$ 为集合 Y 的一个分割。例如 $Y=\{1,2,3\}$,一个分割就是 $Y_1=\{1,2\}$,$Y_2=\{3\}$,而 $Y_1=\{1\}$,$Y_2=\{2\}$,$Y_3=\{3\}$ 则是另一个分割。

小刘会分别计算两个博彩的期望值，并据此来作出行动。这就要求小刘的偏好必须满足 v－N－M 理性偏好的要求，收益函数则为伯努利收益函数。由于动态贝叶斯博弈的最终结果具有不确定性，因而博弈中参与者的偏好实际上就是定义在博彩之上的，因而动态贝叶斯博弈中的偏好同样必须满足 v－N－M 理性偏好的要求。

有了以上的说明，我们就可以给出动态贝叶斯博弈的扩展式定义。

定义 7.1 动态贝叶斯博弈的扩展式一个贝叶斯扩展式博弈为 $\Gamma=\{N,H,P,I,\rho,u\}$，其中① N 为参与者集合，$N=\{0,1,2,\cdots,n\}$，0 代表自然。② H 为全历史集合，即从博弈开始到博弈结束所有可能的行动序列 $\{a^0,a^1,\cdots,a^m\}$，m 为任一自然数，表示一个全历史包含的行动次数。③ P 为参与者函数，即对于每一个子历史 h，$P(h)$ 将其映射成自然或是其他参与者。④ I 为信息空间，I_i 则为参与者 i 的信息空间，每一个信息空间中的信息集必须满足上面提到的三个特性。⑤ ρ 则为自然的概率分布函数，其表示当自然行动时，自然以多大的概率选择某个行动。⑥ u 表示参与者的偏好，其是定义在全历史结果（即博彩）上的 v－N－M收益函数。

前面提到的各种博弈模型实际上都可以用动态贝叶斯博弈的扩展式来表达，策略式博弈、扩展式博弈、贝叶斯博弈都可以看作是一般扩展式的特例。例如，完全信息静态博弈中的囚徒困境的一般扩展式为 $\Gamma=\{N,H,P,I,\rho,u\}$，其中 $N=\{1,2\}$，没有自然；$H=\{CC,CS,SS,SC\}$；$P(\varnothing)=1$，$P(S)=P(C)=2$；$I_1=\{\varnothing\}$，$I_2=\{S,C\}$；由于自然不存在，所以 ρ 也不存在；u 则为收益函数。实际上，如果定义自然只有一个行动，那么 $\rho=1$，也就是说包含自然或不包含自然结果都一样。囚徒困境的博弈树如图 1－7 所示。当囚徒 2 行动时，他不能区分到底位于子历史 S，还是子历史 C（因为他与囚徒 1 同时行动），所以 $\{S,C\}$ 就构成了囚徒 2 的一个信息集。在博弈树中，通常用一条虚线或是方框将相关的节点连起或是圈起。又例如，完全信息动态博弈中的进入博弈的一般扩展式为 $\Gamma=\{N,H,P,I,\rho,u\}$，其中 $N=\{1,2\}$，1 表示挑战者，2 表示在位者，自然不存在；$H=\{(Q),(E,A),(E,F)\}$，Q 表示不进入，E 表示进入，A 表示在位者默许挑战者的进入，而 F 则表示战斗；$P(\varnothing)=1$，$P(E)=2$；$I_1=\{\varnothing\}$，$I_2=\{E\}$；由于自然不存在，所以 ρ 也不存在；u 则为收益函数。进入博弈的博弈树如图4－4所示。由于进入博弈中的每一个信息集都只包含 1 个真子历史，所以进入博弈是完美信息博弈，这就得到一个等价的定义：**如果每一个信息集都只包含一个真子历史，那么博弈就是完美信息博弈。**再例如，贝叶斯信用困境博弈的一般扩展式为 $\Gamma=\{N,H,P,I,\rho,u\}$，其中 $N=\{0,1,2\}$，0 代表自然；$H=\{GHH,GHC,GCH,GCC,BHH,BHC,BCH,BCC\}$，G 和 B 表示自

然的行动,G 表示自然决定厂商 1 为良商,B 则表示自然决定厂商 1 为奸商;厂商 2 是良商,但对厂商 1 是良商还是奸商不清楚;$P(\varnothing)=0$,$P(G)=P(B)=1$,$P(GH)=P(GC)=P(BH)=P(BC)=2$;$I_0=\{\varnothing\}$,$I_{11}=\{G\}$,$I_{12}=\{B\}$,$I_2=\{GH,GC,BH,BC\}$;$\rho=\{1/2,1/2\}$,即自然各以 1/2 的概率选择良商和奸商;$u$ 则为伯努利收益函数。其博弈树如图 7-1 所示,为了简单我们省掉了各全历史下的收益。

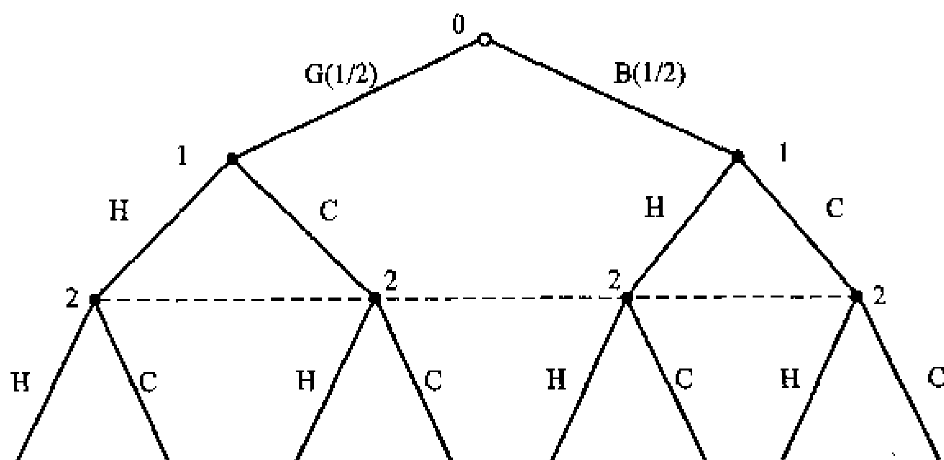


图 7-1　贝叶斯信用困境

【例 7.1】　猜大小

两张牌一张大,一张小,参与者 1 首先随机抽出一张牌。如果他看,那么游戏结束;如果不看,那么下注 1 元钱。然后轮到参与者 2 行动,他可以选择放弃或是跟进。如果放弃,游戏结束;如果跟进,那么也下注 1 元钱,然后揭开牌。如果牌大,那么参与者 1 赢;如果牌小,那么参与者 2 赢。该博弈的一般扩展式为 $\Gamma=\{N,H,P,I,\rho,u\}$,其中:

(1) 参与者集合:$N=\{0,1,2\}$。

(2) 全历史集合:设 B 代表自然选择大,S 代表自然选择小,L 代表参与者 1 选择看,X 代表下注 1 元钱,G 代表参与者 2 放弃,那么全历史集合 $H=\{BL, BXG, BXX, SL, SXG, SXX\}$。

(3) 参与者函数:$P(\varnothing)=0$,$P(B)=P(S)=1$,$P(BX)=P(SX)=2$。

(4) 信息空间:自然的信息集为 $I_0=\{\varnothing\}$,参与者 1 的信息集为 $I_1=\{B, S\}$,即参与者 1 不知道牌是大是小,参与者 2 的信息集为 $I_2=\{BX, SX\}$,同样不知道牌大牌小,但知道参与者 1 的行动。

(5) 自然的概率分布函数:$\rho(B)=1/2$,$\rho(S)=1/2$。

(6) 偏好:如果牌大,两人都下注,那么参与者 1 收益为 1,参与者 2 收益为 -1;如果牌小两人都下注,那么参与者 1 收益为 -1,参与者 2 收益为 1。

该博弈的博弈树如图 7－2 所示。

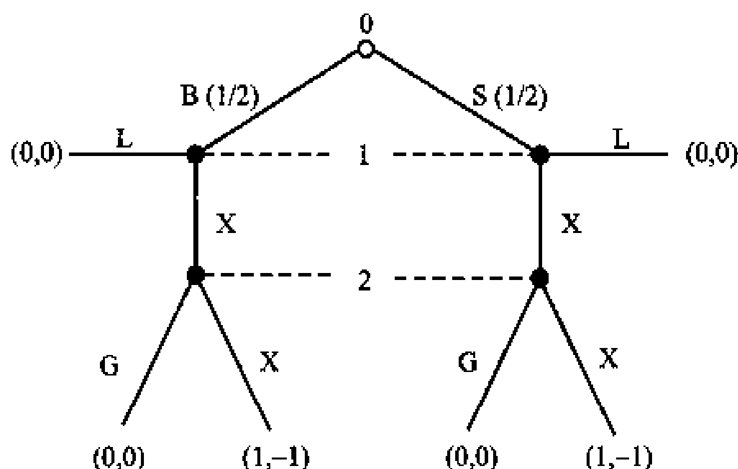


图 7－2 猜大小

【例 7.2】 信号干扰

参与者 1 首先行动，向参与者 2 发出某种信号，信号有两种可能：X 和 Y，但是由于存在干扰，参与者 2 观察到真实信号的概率为 r，而观察到假信号的概率为 $1-r$。例如，参与者 1 发出的信号为 X，但参与者 2 观察到 X 的概率为 r，而有 $1-r$ 的可能观察到 Y。参与者 2 在观察到受干扰的信号后，相应地选择 X 或 Y，博弈结束。对信号的干扰，我们可以理解为自然在接收到参与者 1 的行动 X（或 Y）后，以概率分布 $(r,1-r)$（或 $(1-r,r)$）分别选择行动 X 和 Y。该博弈的一般扩展式为 $\Gamma=\{N,H,P,I,\rho,u\}$，其中：

（1）参与者集合：$N=\{0,1,2\}$。

（2）全历史集合：$H=\{XXX,XXY,XYX,XYY,YXX,YXY,YYX,YYY\}$。

（3）参与者函数：$P(\varnothing)=1, P(X)=P(Y)=0, P(XX)=P(YX)=P(XY)=P(YY)=2$。

（4）信息空间：参与者 1 的信息集为 $I_1=\{\varnothing\}$；自然的信息集为 $I_{01}=\{X\}$，$I_{02}=\{Y\}$；参与者 2 的信息集为 $I_{21}=\{XX,YX\}$，$I_{22}=\{XY,YY\}$，即轮到参与者 2 行动的时候他分不清是历史 XX 还是 YX，对于 $\{XY,YY\}$ 同理。

（5）自然的概率分布函数：$\rho(X|X)=\rho(Y|Y)=r$，即如果参与者 1 选择 X，自然选择 X 的概率为 r。

（6）偏好：具体的收益如图 7－3 所示。

有了一般扩展式的定义，接下来我们将给出一般性的均衡概念——序列均衡。

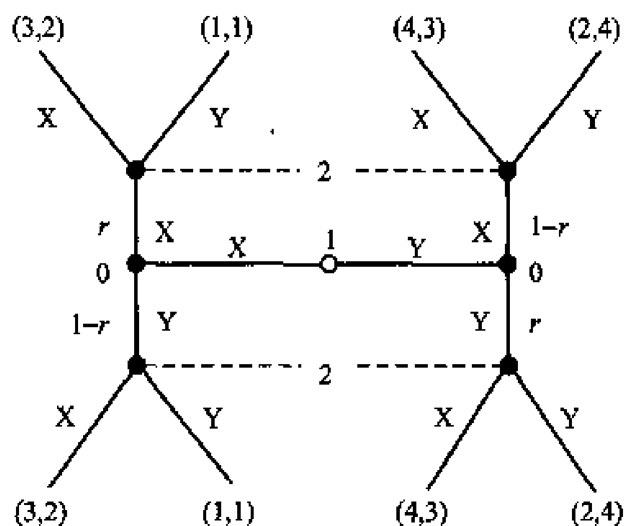


图 7-3 信号干扰

第二节 序列均衡

在这一节中，我们首先从完美贝叶斯均衡入手，然后过渡到序列均衡。完美贝叶斯均衡，实质上是子博弈完美均衡与贝叶斯纳什均衡的综合，它可以分别看作是子博弈完美均衡和贝叶斯均衡的扩展。在非完全信息静态博弈中，我们曾看到参与者如何猜测对手的类型(即信号)是相当关键的，参与者的策略实际上是类型依存性的。类型猜测(即信念)不同得到的策略也会不同，这说明参与者对类型的猜测与他的策略具有同样的重要性，而在非完全信息动态博弈中，这一思想得到了进一步的加强。非完全信息动态博弈中的均衡概念不再是策略组合，而是策略+信念的组合。克瑞普斯和威尔逊的贡献，就在于将信念与策略结合了起来，使得人们能够很方便地分析动态博弈中的信息不完全问题。运用子博弈完美均衡的方法，同时加上对猜测的合理设定，从而解决了在不完全信息下如何设定猜测、如何定义均衡、如何计算均衡这些问题。

一、完美贝叶斯均衡(perfect bayesian equilibrium)

所谓完美是指在动态博弈中，整体的纳什均衡也必须是所有子博弈的纳什均衡，但当一个信息集在某些策略下走到的可能性为0时，我们实际上没有办法对这个策略是否在这个信息集中仍是最优反应加以验证，因为子博弈完美均衡没有提供在可能性为0下我们如何行动的指南。具体讲，当某一个信息集在整体策略走到的概率为0、且这个信息集包含多个子历史时，我们就没有办法运用贝叶斯法则去估计各子历史的概率是多少。在这种状况下，由于参与者无法计

算他们的收益,因此子博弈完美均衡会忽略掉这个信息集,而不去检查整体策略在这个信息集及其后面的所有信息集中是否为最优反应,一旦如此就会产生很不合理的结果。这里我们要定义一个新的子博弈概念,即泛子博弈。泛子博弈与子博弈的唯一区别就是泛子博弈的起始点为一个信息集,而不论这个信息集是否只包含一个真子历史。

定义 7.2 一个动态贝叶斯扩展式博弈 $\Gamma = \{N,H,P,I,\rho,u\}$ 中的**泛子博弈**为 $\Gamma(I)$,是该博弈的一个特殊子集,其具有下列特征:

(1) 它必须从一个信息集开始,并包含该信息集其后所有的历史。

(2) 泛子博弈不能"分割"任何信息集。

现在我们通过图 7-4 所示博弈来说明子博弈完美均衡存在的问题。

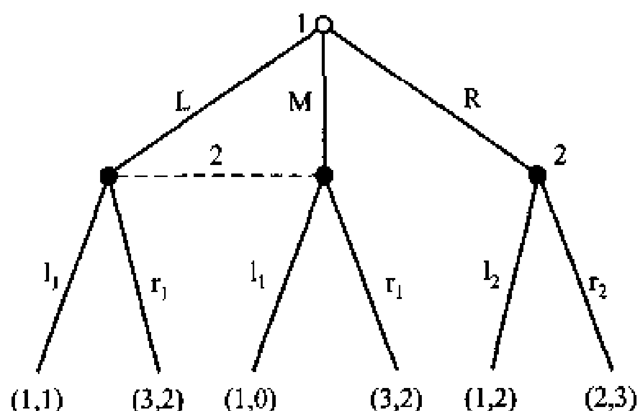


图 7-4 动态贝叶斯博弈示例

如图 7-4 所示的博弈存在一个子博弈完美均衡:参与者 1 选择 R,参与者 2 在信息集{R}选择 r_2,在信息集{L,M}选择 l_1。证明的方法如下:给定参与者 2 的策略,若参与者 1 选择 M 的话,他的报酬是 1;若参与者 1 选择 L,他的报酬也为 1;而如果参与者 1 选择 R,那么收益为 2,所以参与者 1 选择 R 是最优反应。反之,在参与者 1 的策略为 R 的情况下,参与者 2 在信息集{R}选择 r_2,在信息集{L,M}选择 l_1 的策略是最优反应。因为在参与者 1 的策略为 R 的情况下,{L,M}这个信息集不可能会走到,因此参与者 2 在信息集{L,M}上无论策略是什么都不会影响到他所获得的收益,而且从{L,M}出发的这个"博弈"不是一个子博弈(而是泛子博弈)。这表明策略组合:参与者 1 选择 R,参与者 2 在信息集{R}选择 r_2,在信息集{L,M}选择 l_1,既是该博弈的纳什均衡,也是该博弈的子博弈完美均衡。

但仔细观察后发现,如果博弈真的进行到了信息集{L,M},参与者 2 真的会选择 l_1 吗?答案是否定的,因为当参与者 2 走到{L,M}时,他选择 r_1 的收益等于 2,而他选择 l_1 则等于 1 或 0,因此不论他在子历史 L 还是 M,选择 r_1

一定比 l_1 要好。换言之,在信息集{L,M}下,选择 l_1 其实是一个严格劣行动。出现这种问题的原因在于,当参与者 1 的策略为 R 时,信息集{L,M}不会发生,因而参与者 2 在信息集{L,M}下无论如何行动都不会影响他的收益,这就造成子博弈完美均衡中有可能包含着违反理性选择的情况。这也说明,在动态贝叶斯博弈中光有子博弈完美均衡是远远不够的,我们需要发展出新的均衡概念。

在完全信息动态博弈中,我们反复强调子博弈完美均衡最大的优点就是可以排除所谓的不可置信威胁。如果我们将这一思想扩展到泛子博弈(而不仅仅是子博弈),那么能不能克服如图 7-4 中出现的问题呢?要将子博弈完美均衡的概念运用到泛子博弈,首先需要解决的问题就是参与者如何作出猜测,这时参与者所谓完整的行动计划不再仅仅包含策略,还应包含对所有信息集概率的猜测,而不管这个信息集能否走到,这个猜测称为信念系统。

定义 7.3 动态贝叶斯博弈中的一个**信念系统**(belief system)是一个概率函数,其给出每一个信息集中所含子历史的一个概率分布(不管这个信息集能否走到),即一个信息集所含各子历史的概率大于等于 0,相加等于 1。

那么在动态贝叶斯博弈下,策略是如何定义的呢?

定义 7.4 动态贝叶斯博弈中的参与者 i 的一个**行为策略**(behavioral strategy)就是一个函数,其给出参与者 i 在每一个信息集 I_i 中的行动 $A(I_i)$ 的概率分布,并且各概率分布在各信息集中彼此独立。

例如,在图 7-4 中,参与者 2 的一个信念系统就是当博弈进行到信息集{L,M}时,参与者 2 猜测子历史 L 发生的概率为 r,而子历史 M 发生的概率为 1-r。参与者 1 在信息集{∅}下的一个行为策略就是以 θ_1、θ_2、$1-\theta_1-\theta_2$ 的概率分别选择 L,M 和 R。参与者 2 的一个行为策略就是在信息集{L,M}下,以 $1\geq p\geq 0$的概率选择 l_1,$1-p$ 的概率选择 r_1;在信息集{R}下,以 $1\geq q\geq 0$ 的概率选择 l_2,$1-q$ 的概率选择 r_2。

所谓完美贝叶斯均衡,其核心就在于它要求即使参与者碰到猜测上的困难,参与者仍应该想办法"制造"出一个对信息集内的各子历史的可能猜测。一个"好"的均衡,不但要参与者说明在何处该做什么决定,而且参与者的心中也时时要有一个自己在各个子历史的概率猜测。当博弈进行到任意一个历史时,每一个参与者早就胸有成竹——猜测出不同历史的概率各是多少。克瑞普斯和威尔逊的贡献之一就是规定了参与者应当如何"制造"对这些概率的猜测——即应该根据博弈原先的设定和参与者所用的策略去制造猜测。即使有些历史完全不可能走到,参与者也要想办法去设定猜测。因此在每一个参与者的心中,不但要有一个策略存在,还要有一个信念系统,这个信念系统告诉他不同历史的概率

各是多少。这样一个策略加信念的东西，就是参与者的审时度势(assessment)。一个审时度势实质上就是参与者的万全之策。

定义 7.5 动态贝叶斯博弈中参与者 i 的一个**审时度势**(assessment)是他的行为策略加上信念系统，记为(σ_i,μ_i)。其中 σ_i 是他的行为策略，而 μ_i 是一个定义在每一个信息集 $I_{ik}\in I_i$ 之上的概率分布函数。也就是说，对每一个 $I_{ik}\in I_i$ 和 $h\in I_{ik}$；$\mu_i(h)\geqslant 0$，且 $\sum_{h\in I_{ik}}\mu_i(h)=1$。在任何一个信息集 I_{ik}上，μ_i 所代表的意思是，如果博弈的历史进行到 I_{ik}，那么参与者 i 认为他在其中某个子历史 $h\in I_{ik}$ 的概率为 $\mu_i(h)$。

需要强调的是，参与者的猜测和策略就像一个人的两只手，相互协调实现共同的目标。所谓最优策略是在给定猜测条件下的最优，所谓猜测是在最优策略下的猜测。更精确地说，猜测必须由策略经过贝叶斯法则计算出来，而策略必须是给定猜测下的最优反应。假设 I 是某一信息集，而(σ,μ)是一个审时度势，且 $h\in I$，那么当博弈进行到 I 这个信息集时，在(σ,μ)这个审时度势下，博弈现在正位于历史 h 的概率为

$$\mu(h)=\frac{\pi^{\sigma,\rho}(h)}{\pi^{\sigma,\rho}(I)} \tag{7-1}$$

其中 $\pi^{\sigma,\rho}(I)=\sum_{h\in I}\pi^{\sigma,\rho}(h)$。上式实际上就是贝叶斯法则，式中的分母表示在先验猜测 ρ(自然行动的概率)和策略组合 σ 之下，博弈进行到信息集 I 的概率。而分子则代表博弈进行到历史 h 的概率。因此当参与者身处 I 时，他认为自己在 h 的概率 $\mu(h)$，就等于 $\pi^{\sigma,\rho}(h)/\pi^{\sigma,\rho}(I)$。但是当 I 走到的概率为 0，也就是说 $\pi^{\sigma,\rho}(I)=0$ 时，由于分母为 0，贝叶斯法则无法使用，因此 $\mu(h)$无定义。这时对历史 h 的猜测应如何去设定呢？一个最简单的做法，是随便去设定，但所用的策略必须是这个猜测的最优反应。

定义 7.6 在动态贝叶斯扩展式博弈 $\Gamma=\{N,H,P,I,\rho,u\}$ 中，若一个审时度势组合(σ^*,μ^*)满足：

(1) 若某个信息集 I 走到的概率大于 0，那么 μ^* 在 I 上的值一定是从 σ^* 经由贝叶斯法则计算出来的；

(2) 若某个信息集 I 走到的概率为 0，那么 μ^* 在 I 上的值无法从 σ^* 经贝叶斯法则计算出来，则 μ^* 在 I 上的值可以任意设定；

(3) 对每一参与者 i 的任一行动策略 σ_i，不等式

$$u_i(\sigma_i^*,\sigma_{-i}^*,\mu^*\mid\Gamma(I))\geqslant u_i(\sigma_i,\sigma_{-i}^*,\mu^*\mid\Gamma(I))$$

对每一个泛子博弈 $\Gamma(I)$ 成立。

则我们称(σ^*,μ^*)为一个**完美贝叶斯均衡**(perfect bayesian equilibrium)。

完美贝叶斯均衡所要求的条件,其实相当简单。首先,它要求参与者的信念系统必须从策略出发"理性地"加以计算。如果算不出(也就是某个非单点信息集 I 走到的可能性为0,即 $\pi^{\sigma^*,\mu^*}(I)=0$),那么信念可以随意去设。换言之,这时对任何 $h\in I$,$\mu^*(h)$ 的值可以任意猜,只要 $\sum_{h\in I}\mu^*(h)=1$ 就可以。其次,它要求任一参与者所用的策略,必须是每一个泛子博弈(不只是子博弈),相对于 μ^* 这个信念系统的纳什均衡。

我们现在要证明,利用完美贝叶斯均衡的定义,图7-4中不合理的子博弈完美均衡[R;(l_1,r_2)]并不是一个真正的均衡——完美贝叶斯均衡。

假设(σ^*,μ^*)是一个完美贝叶斯均衡,其中 σ^* 是行动策略组合[R;(l_1,r_2)]。很明显,所有参与者对信息集{∅}和{R}的猜测都为1,即 $\mu_1^*(\varnothing)=\mu_2^*(\varnothing)=1$,且 $\mu_1^*(R)=\mu_2^*(R)=1$。我们唯一要算的,是在信息集{L,M},$\mu_2^*(L)$ 和 $\mu_2^*(M)$ 各应是多少。由于在策略 σ^* 下,{L,M}走到的可能性为0,所以 μ_2^* 在{L,M}上的值,无法经由贝叶斯法则从 σ^* 计算出来。根据完美贝叶斯均衡的定义,$\mu_2^*(L)$ 和 $\mu_2^*(M)$ 的猜测可随意设定。不妨令 $\mu_2(L)=t$,因而任何一个 $1\geqslant t\geqslant 0$ 都符合要求。在给定这个猜测下,参与者2选择 l_1 的预期收益为

$$u_2(\sigma^*,\mu^*\mid\{L,M\})=t\cdot 1+(1-t)\cdot 0=t$$

选择 r_1 的预期收益为

$$u_2(\sigma^*,\mu^*\mid\{L,M\})=t\cdot 2+(1-t)\cdot 2=2。$$

即无论 t 的值是多少,2都大于 t,因此 l_1 绝对不会是参与者2的最优反应。换言之,参与者2在信念 μ^* 下的策略 $\sigma_2^*(\{L,M\})=l_1$,并不是泛子博弈{L,M}的纳什均衡,它与定义7.6中的条件(3)相矛盾。所以不论对历史L的信念 $\mu(L)$ 是多少,σ^* 都不可能是一个完美贝叶斯均衡。之所以会有这样的结果,原因非常明显。在这个例子中,在行为策略组合 σ^* 下,信息集{L,M}走到的可能性为0,从而参与者2无法根据贝叶斯法则计算他在历史L和M的概率各是多少。而根据于博弈完美均衡的定义,我们无须检查参与者2的策略在信息集{L,M}中是否为最优反应。但完美贝叶斯均衡的要求则进了一步,它要求不管信息集{L,M}能否走到,参与者一定要对L和M的可能性做一个猜测。更重要的是,一旦有了这个(不可能走到的信息集的)猜测,他的策略一定要是这个猜测的最优反应,即他的信念一定与策略完全一致。而在图7-4中,由于 l_1 是参与者2的严格劣行动,因此不论 $\mu_2(L)$ 或 $\mu_2(M)$ 是多少,他一定会(理性地)选择 r_1,这才是对应于 $\mu_2(L)$ 的最优反应。也正因如此,图7-4中的子博弈完美均衡 σ^* 并不是一个完美贝叶斯均衡。

那么这个博弈的完美贝叶斯均衡是什么？我们可以用逆推法计算。很明显，在{R}这个泛子博弈，参与者 2 应该选择 r_2。在泛子博弈{L,M}，假设参与者 2 猜测 L 的概率为 $\mu_2(L)=t$，而在 M 的概率为 $\mu_2(M)=1-t$，其中 $1\geq t\geq 0$。若参与者 2 选择 l_1，他的收益为 $t+0(1-t)=t$，若他选择 r_1，那么他的收益为 $2t+2(1-t)=2$。由于 $2>t$，所以参与者 2 一定会选择 r_1。因此不论 t 值为多少，参与者 2 都一定会选择 r_1。给定参与者 2 如上的策略，我们再往前看参与者 1 在{∅}时，他应怎么做。如果他选择 L，那么由于他知道参与者 2 接着会选择 r_1，此时他的收益为 3。同理如果他选择 M，则他的收益也为 3。如果他选择 R，则他的收益为 2。因此参与者 1 随机选择 L 和 M 是对参与者 2 策略的最优反应。这意味着图 7－4 所示的博弈存在许多的完美贝叶斯均衡：

$\sigma_1(\{\varnothing\})=(sL,(1-s)M)$①，其中 $1\geq s\geq 0$；$\sigma_2(\{L,M\})=r_1$，$\sigma_2(\{R\})=r_2$。相对应的信念为 $\mu_1(\{\varnothing\})=\mu_2(\{\varnothing\})=1$，$\mu_1(R)=\mu_2(R)=1$，$\mu_2(L)=t$，$\mu_2(M)=1-t$，其中 $1\geq t\geq 0$。

二、序列均衡(sequential equilibrium)

序列均衡的要求比完美贝叶斯均衡更进一步，是对完美贝叶斯均衡的再精炼。它要求即使某个信息集在策略下走到的概率为 0，参与者的猜测也不能任意设定，因而序列均衡比起完美贝叶斯均衡要更合理。在正常的情况下，有些信息集走到的概率为 0，但人非圣贤，孰能无过，一旦参与者无意中出错，走到了原来策略不可能走到的信息集，参与者会随意去设定他的猜测吗？显然这样的要求是不合理的。如果一个人端着一杯滚开水，突然手发抖，滚开水泼了一身，发生了烫伤，那怎么办，难道是随意处理吗？这种考虑参与者会犯错误、并在错误发生之后应当如何处理的思想就是序列均衡的核心。具体讲，假设某一个参与者有三个可用策略，L、M 和 R。如果他决定选择 R，而不选择其他策略，那么其他参与者仍然会认为他有可能犯错而选择 L 或 M。虽然他选择 L 或 M 的概率很小，但永远大于 0。换言之，选择 R 的这个策略等价于 $(\varepsilon_1L,\varepsilon_2M,(1-\varepsilon_1-\varepsilon_2)R)$ 这个策略，其中 ε_1 和 ε_2 无限趋近于 0。这样做的意义在于，无论是什么样的策略，任何信息集走到的概率一定为正，这就保证了(7－1)式的分母永远不为 0，因而参与者的信念永远可根据贝叶斯法则从参与者的策略中估算出来。

定义 7.7 在动态贝叶斯扩展式博弈 $\Gamma=\{N,H,P,I,\rho,u\}$ 中，若一个审时

① 表示以 s 的概率选择 L，以 $1-s$ 的概率选择 M。

度势(σ^*,μ^*)满足下列条件，则称它是一个**序列均衡**：

（1）对每一参与者 i 的任一行动策略 σ_i，不等式

$$u_i(\sigma_i^*,\sigma_{-i}^*,\mu^* \mid \Gamma(I)) \geqslant u_i(\sigma_i,\sigma_{-i}^*,\mu^* \mid \Gamma(I))$$

对每一个泛子博弈 $\Gamma(I)$ 成立。

（2）存在一个序列 $\{(\sigma^t,\mu^t)\}_{t=1}^{\infty}$，使得 $(\sigma^*,\mu^*) = \lim_{t\to\infty}(\sigma^t,\mu^t)$，其中每一个 σ^t 都是行为策略组合，且 μ^t 是由 σ^t 经贝叶斯法则计算得来。也就是说，每一个 μ^t 和 σ^t 之间，都必须满足(7-1)式[①]。

定义7.7的第一条被称为**序列理性**（也称序贯理性 sequential rationality），实际上是对子博弈完美均衡要求的推广，第二条被称为**信念与策略的一致性**（consistency of beliefs with strategies），它表明信念与策略是你中有我、我中有你的，共同被决定。从这个定义还可看出，所谓序列均衡，其实是一个策略序列的极限，即它假设所有参与者使用任何行为策略的概率都为正，尽管可能其被使用的概率趋近于0，并且这个策略的极限必须满足理性的要求（即定义7.7中的(1)式）。需要注意的是，我们只要求(σ^*,μ^*)（无穷序列$\{(\sigma^t,\mu^t)\}_{t=1}^{\infty}$的极限）满足定义7.7中的(1)式，而没有要求序列(σ^t,μ^t)也必须满足定义7.7中的(1)式。这意味着对$\{(\sigma^t,\mu^t)\}_{t=1}^{\infty}$这个序列的选择弹性非常大，这构成了序列均衡的一个缺点，有了后来学者对序列均衡的再精炼。通过比较，也可看出完美贝叶斯均衡和序列均衡最大的差异就是对参与者信念的设定，序列均衡在信念上的设定比完美贝叶斯均衡合理，但计算要繁杂得多，正所谓有所得必有所失。不过幸运的是，在绝大多数情况下，完美贝氏均衡和序列均衡完全一样，两者不同倒是一个特例。

在如图7-5所示的博弈中，自然（nature）选择 x 和 y 的概率各为1/3和2/3，即 $\rho(x)=1/3$。对于自然的选择参与者1和参与者2都不能观察到。参与者1可以选择L或R，参与者2能观察到参与者1的行动，但不知道参与者1是从历史x和y中哪一点出发的。不妨假设存在这样一个均衡策略 σ^*，其中 $\sigma_1^*(L)=1$（表示参与者1以概率1选择L），这意味着 $\pi^{\sigma,\rho}(\{xR,yR\})=0$。在完美贝叶斯均衡的定义下，由于$\{xR,yR\}$这个信息集走到的概率为0，所以参与者2在信息集$\{xR,yR\}$的信念 $\mu_2(xR)$（或 $\mu_2(yR)$）可以是0到1之间的任意数，只要参与者2在信息集$\{xR,yR\}$的策略是相对于 $\mu_2(xR)$ 和 $\mu_2(yR)$ 的最优反应即可。

① 序列均衡只要求 σ_i^* 是相对于 σ_{-i}^* 的最优反应。而不要求 σ_i^t 是 σ_{-i}^t 的最优反应。换言之，(σ^t,μ^t) 不需要满足定义7.7中的(1)式。

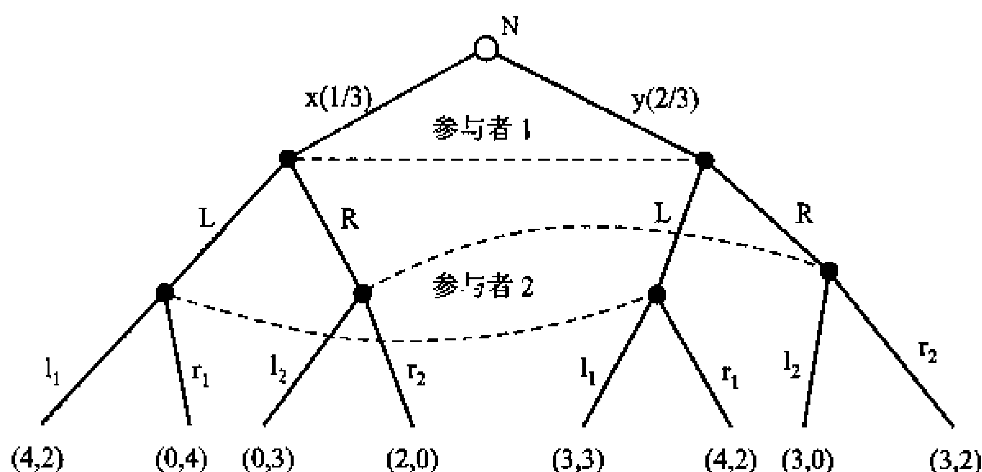


图 7-5 非完全信息博弈

例如，不妨设 $\mu_2(xR)=t$，那么参与者 2 选择 l_2 的期望值为 $3t+0\times(1-t)=3t$，选择 r_2 的期望值为 $0\times t+2(1-t)=2-2t$，如果参与者 2 选择 l_2 优于选择 r_2，那么要求 $1\geqslant t\geqslant 2/5$（如果参与者 2 选择 r_2 优于选择 l_2，那么 $0\leqslant t\leqslant 2/5$）。但在序列均衡的定义下，就没有这么大的自由度。不妨假设 $\sigma_1^t(L)=1-\varepsilon^t$，$\sigma_1^t(R)=\varepsilon^t$，$1>\varepsilon^t>0$，对于 $\forall t$，且 $\lim\limits_{t\to\infty}\varepsilon^t=0$。如此一来，$\sigma_1^*$ 就是 σ_1^t 的极限。在这个设定下，

$$\mu_2^t(xR)=\frac{1/3\varepsilon^t}{1/3\varepsilon^t+2/3\varepsilon^t}=1/3$$

对任意 t 成立。因此 $\mu_2(xR)=\lim\limits_{t\to\infty}\mu_2^t(xR)=1/3$，即在序列均衡的设定下，如果 $\sigma_1^*(L)=1$，那么 $\mu_2^*(xR)$ 非是 1/3 不可。这个例子说明，当某一信息集在策略组合 σ^* 下走到的概率为 0 时，完美贝叶斯均衡可以随意设定信念，而序列均衡则必须尊重原先博弈的信息结构，在参与者信念的设定上，做了“如果出错，仍然理性”这样的限制。但也因如此，造成了求解序列均衡较为繁琐的结果。所幸在大多数情况下，两者是一样的，不一样反而是特例。我们用图 7-6 再加以说明。

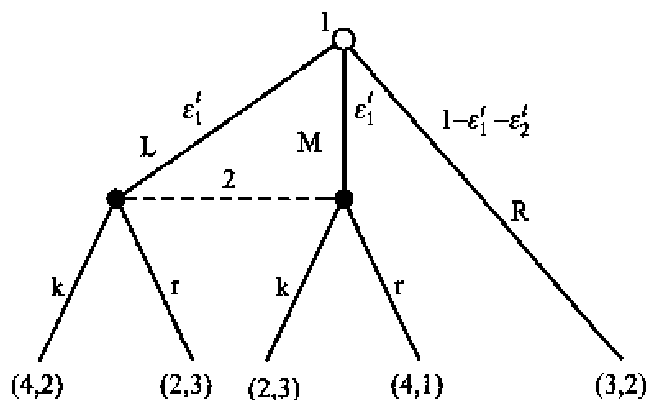


图 7-6 一个例子(1)

假设如图7－6所示的动态博弈存在一个均衡策略，其中 $\sigma_1^*(\{\varnothing\})=R$，因此 $\pi^{\sigma,\rho}(\{L,M\})=0$。根据完美贝叶斯均衡的定义，$\mu_2(L)$ 可以为任一个0到1之间的数，而现在我们则根据序列均衡的定义，看一看对 μ_2 有些什么限制。令 $\sigma_1^t(R)=1-\varepsilon_1^t-\varepsilon_2^t$，$\sigma_1^t(L)=\varepsilon_1^t$，$\sigma_1^t(M)=\varepsilon_2^t$，且 $\lim_{t\to\infty}\varepsilon_1^t=\lim_{t\to\infty}\varepsilon_2^t=0$，则 $\lim_{t\to\infty}\sigma_1^t=\sigma_1^*$，并且

$$\mu_2^t(L)=\frac{\varepsilon_1^t}{\varepsilon_2^t+\varepsilon_1^t};\quad \mu_2^t(M)=\frac{\varepsilon_2^t}{\varepsilon_2^t+\varepsilon_1^t}$$

对任意一个 $1\geqslant s\geqslant 0$，令 $\varepsilon_1^t=\frac{s}{1-s}\varepsilon_2^t$，那么 $\mu_2^t(L)=s$，$\mu_2^t(M)=1-s$。这说明，该例子中的序列均衡与完美贝叶斯均衡是完全一样的。原因在于，虽然序列均衡要求用一个无穷序列去逼近均衡策略，但是无穷序列本身的选择却非常随意，这种随意与完美贝叶斯均衡中的随意设定猜测在大多数情况下是完全一样的。因而我们说序列均衡与完美贝叶斯均衡相同是常例，不相同倒是特例。这一点可以帮助我们最大程度地简化对序列均衡的计算。下面我们仍用图7－6来说明序列均衡的算法。

首先，通过观察我们发现，如果参与者1选择L，那么参与者2会选择r；如果参与者2选择r，那么参与者1更愿意选择M（而不是L）；如果参与者1选择M，那么参与者2就会选择k（而不是r）。这说明对于参与者1而言，存在着两种可能：一是选择R，游戏就此结束；一是在L和M之间随机选择，而不选择R。

对于第一种情况，参与者1选择R，那么信息集{L,M}就不可能走到。通过前而的介绍，我们已经知道该博弈的序列均衡等价于完美贝叶斯均衡。不妨设参与者2在信息集{L,M}的信念为 $\mu_2(L)=s$，$1\geqslant s\geqslant 0$，那么参与者2选择k的预期收益为 $2\mu_2(L)+3\mu_2(M)=3-s$，选择 r 的预期收益为 $3\mu_2(L)+1\mu_2(M)=1+2s$，所以当 $s<2/3$ 时，参与者2的最优反应是k；当 $s>2/3$ 时，最优反应是r；当 $s=2/3$ 时，他的最优反应为概率分布 $(p,(1-p))$（即以 p 的概率选择k），$1\geqslant p\geqslant 0$。

对于第二种情况，参与者1在L和M之间随机选择，而不选择R。这就要求L和M的预期收益必须相等，并且还必须大于等于R的收益。不妨设参与者2的行为策略为概率分布 $(p,(1-p))$，那么参与者1选择L的期望收益为 $4p+2(1-p)$，选择M的期望收益为 $2p+4(1-p)$，要使得L和M的预期收益相等，就必须要求 $p=1/2$；同时，选择L（或M）的收益还不能小于选择R的收益，即 $4p+2(1-p)\geqslant 3$（或 $2p+4(1-p)\geqslant 3$），这同样要求 $p=1/2$。因为，当 $p<1/2$

时，L 小于 R 的收益，但 M 大于 R 的收益；当 $p>1/2$ 时，M 小于 R 的收益，但 L 大于 R 的收益；当 $p=1/2$ 时，L，M，R 的收益都等于 3。反过来，要使参与者 2 随机选择 k 和 r，必须使得 l 和 r 的预期收益相等，不妨设参与者 1 选择 L 的概率为 s，选择 M 的概率为 $1-s$，那么参与者 2 选择 k 的预期收益为 $2s+3(1-s)$，选择 r 的预期收益为 $3s+(1-s)$，解出的结果为 $s=2/3$。综上所述，该博弈的所有序列均衡为：

(1) $\sigma_1(\{\varnothing\})=(0,0,1)$（即选择 R），当 $\mu_2(\mathrm{L})<2/3$ 时，$\sigma_2(\{\mathrm{L,M}\})=\mathrm{k}$；当 $\mu_2(\mathrm{L})>2/3$ 时，$\sigma_2(\{\mathrm{L,M}\})=r$；当 $\mu_2(\mathrm{L})=2/3$ 时，$\sigma_2(\{\mathrm{L,M}\})=(p,1-p)$，$1\geqslant p\geqslant 0$。

(2) $\sigma_1(\{\varnothing\})=(2/3,1/3,0)$（即随机选择 L 和 M），$\mu_2(\mathrm{L})=2/3$，$\sigma_2(\{\mathrm{L,M}\})=(1/2,1/2)$。

在第一种情况下，读者可能会问，如果 $\sigma_2(\{\mathrm{L,M}\})=\mathrm{k}$，那么参与者 1 选择 L（收益为 4）将好过选择 R（收益为 3），但问题是若参与者 1 选择 L，那么 $\mu_2(\mathrm{L})=1$，而不可能小于 2/3（这表明信念和策略不一致）。同理，如果 $\sigma_2(\{\mathrm{L,M}\})=r$，那么参与者 1 的最优反应是选择 M，但若如此，$\mu_2(\mathrm{L})$ 就应该等于 0，而不是大于 2/3。这充分说明，序列均衡（或完美贝叶斯均衡）下信念和策略必须一致，所谓的均衡必须由信念 + 策略所组成。

下面我们再考虑对图 7－6 的一个变形，见图 7－7 所示博弈。

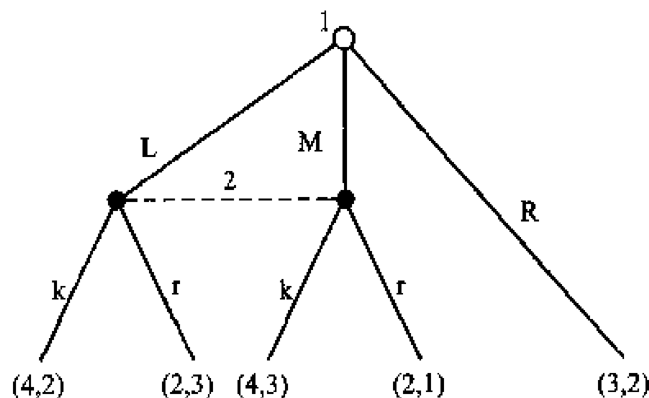


图 7－7　一个例子(2)

我们运用先逆后顺的方法。假设参与者 2 在 $\{\mathrm{L,M}\}$ 的信念为 $\mu_2(\mathrm{L})=t$，$\mu_2(\mathrm{M})=1-t$。则他选择 k 时获得的预期收益是 $2t+3(1-t)=3-t$，而选择 r 的预期收益是 $3t+1-t=1+2t$。因此，当 $t<2/3$ 时，参与者 2 会选择 k，而 $t>2/3$ 时，参与者 2 会选择 r。$t=2/3$ 时，由于 k 和 r 两者预期收益相等，参与者 2 会选择混合策略 $(v,(1-v))$，$1\geqslant v\geqslant 0$。我们将讨论分成两部分，而其重点在于参与者 1 是否会用 100% 的概率选择 R。因为只有参与者 1 用 100% 的概率选

择 R 时，才需要参与者 2 在信息集{L,M}做猜测，在其他情况下只要用贝叶斯法则去计算就可以了。

(1) 参与者 1 假如未用 100% 的概率选择 R，则假设参与者 1 在空历史{∅}上的策略是 $\sigma_1(\{\varnothing\}) = (s_1, s_2, (1 - s_1 - s_2))$，其中 $1 > s_1, s_2 > 0, 1 > s_1 + s_2 > 0$。

① 若 $s_1 < 2s_2$，则 $\mu_2(L) < 2/3$，参与者 2 选择 k。这时参与者 1 的收益是 $4s_1 + 4s_2 + 3(1 - s_1 - s_2) = 3 + s_1 + s_2$。因此让 $s_1 + s_2 = 1$ 会使参与者 1 的收益最大，等于 4。因此一个均衡是：$\sigma_1(\{\varnothing\}) = (s, (1 - s), 0), s < 2/3, \mu_2(\{L\}) < 2/3$，$\sigma_2(\{L, M\}) = k$。

② 若 $s_1 > 2s_2$，则 $\mu_2(L) > 2/3$，参与者 2 会选择 r。这时参与者 1 的收益为 $2s_1 + 2s_2 + 3(1 - s_1 - s_2) = 1 - (s_1 + s_2)$。当 $s_1 + s_2 = 0$ 时，参与者 1 的收益最大。但这意味着参与者用 100% 的概率选择 R，和我们的假设不符，因此 $s_1 > 2s_2$ 时没有均衡。

③ 若 $s_1 = 2s_2$，则参与者 2 的策略为 $(v, (1 - v)), 1 \geqslant v \geqslant 0$，而参与者 1 的预期收益是 $s_1[4v + 2(1 - v)] + s_2[4v + 2(1 - v)] + 3(1 - s_1 - s_2) = 3 + 3s_2\ (2v - 1)$。若 $v > 1/2$，则 s_2 越大，参与者 1 的收益越高。但由于 $s_1 = 2s_2$，所以 $s_1 = 2/3, s_2 = 1/3$ 时，参与者 1 的收益最大。因此有一个均衡：$\sigma_1(\{\varnothing\}) = (2/3, 1/3, 0), \mu_2(L) = 2/3, \mu_2(M) = 1/3, \sigma_2(\{L, M\}) = (v, (1 - v)), 1 > v > 1/2$。若 $v < 1/2$，则 $s_1 = s_2 = 0$ 时，参与者 1 的收益最大，但这又和参与者 1 不用 100% 的概率下 R 的假设不符，因此，$v < 1/2$ 没有均衡存在。若 $v = 1/2$，无论 s_2 等于多少，参与者 1 的收益都是 3，这个情形下的均衡是：$\sigma_1(\{\varnothing\}) = (2s, s, (1 - 3s)), 0 \leqslant s \leqslant 1/3$，$\mu_2(L) = 2/3, \mu_2(M) = 1/3, \sigma_1(\{L, M\}) = (1/2, 1/2)$。

(2) 假如参与者 1 以 100% 的概率选择 R，那么参与者 2 便需要预测自己在信息集{L,M}下会作什么决定。当 $t < 2/3$ 时，参与者 2 应选择 k。但在 $t < 2/3$ 时，参与者 1 背离到 M 的收益为 4 会大于下 R 的收益 3（这表明 $t = 0$），因此选择 R 不是参与者 1 的最优反应。当 $t > 2/3$ 时，参与者 2 会选择 r，因此参与者 1 选择 R 的收益 3 大于选择 L 或 M 的收益 2。所以 R 是参与者 1 的最优反应。这表明：$\sigma_1(\{\varnothing\}) = R, \sigma_2(\{L, M\}) = r, \mu_2(L) > 2/3$，且 $\mu_2(M) = 1 - \mu_2(L)$ 是一组序列均衡。当 $t = 2/3$ 时，参与者 2 会采用混合策略 $(v, (1 - v))$，因为参与者 1 选择 L 的预期收益等于选择 M 的预期收益，为 $4v + 2(1 - v) = 2 + 2v$。由于选择 R 必须是参与者 1 的最优反应，因此 $2 + 2v \leqslant 3$，即 $v \leqslant 1/2$。因此 $\sigma_1(\{\varnothing\}) = R, \sigma_2(\{L, M\}) = (v, (1 - v)), 1/2 \geqslant v \geqslant 0, \mu(L) = 2/3, \mu(M) = 1/3$，也是一组序列均衡。

为了便于理解，我们将前面所有可能出现的情况列成表 7－1。通过表 7－1 可以很容易地找出图 7－7 所示博弈的全部序列均衡。

表 7-1 序列均衡计算表

当 $s_1>0, s_2>0, 1-s_1-s_2>0$ 时				
$s_1<2s_2$	$s_1>2s_2$	$s_1=2s_2$		
		$v>1/2$	$v=1/2$	$v<1/2$
$\sigma_1(\{\varnothing\})$ $=(s,(1-s),0)$, $s<2/3$, $\mu_2(\{L\})<2/3$, $\sigma_2(\{L,M\})=k$	没有均衡	$\sigma_1(\{\varnothing\})$ $=(2/3,1/3,0)$ $\mu_2(L)=2/3$, $\mu_2(M)=1/3$, $\sigma_2(\{L,M\})$ $=(v,(1-v))$	$\sigma_1(\{\varnothing\})$ $=(2s,s,(1-3s))$, $0\leqslant s\leqslant 1/3$, $\mu_2(L)=2/3$, $\mu_2(M)=1/3$, $\sigma_2(\{L,M\})$ $=(1/2,1/2)$	没有均衡

当 $s_1=s_2=0$，参与者 1 以 100% 的概率选择 R 时

$\mu_2(L)<2/3$	$\mu_2(L)=2/3$	$\mu_2(L)>2/3$
矛盾	$\sigma_1(\{\varnothing\})=R$, $\sigma_2(\{L,M\})=(v,(1-v))$, $1/2\geqslant v\geqslant 0$, $\mu_2(L)=2/3, \mu_2(M)=1/3$,	$\sigma_1(\{\varnothing\})=R$, $\sigma_2(\{L,M\})=r$, $\mu_2(L)>2/3$, $\mu_2(M)=1-\mu_2(L)$

根据表 7-1，这个博弈所有的序列均衡为：

(1) $\sigma_1(\{\varnothing\})=(s,(1-s),0), s\leqslant 2/3, \mu_2(L)\leqslant 2/3, \sigma_2(\{L,M\})=(v,(1-v)), 1\geqslant v\geqslant 1/2$；

(2) $\sigma_1(\{\varnothing\})=R, \sigma_2(\{L,M\})=r, \mu_2(L)>2/3$；

(3) $\sigma_1(\{\varnothing\})=R, \sigma_2(\{L,M\})=(v,(1-v)), 1/2\geqslant v\geqslant 0, \mu_2(L)=2/3$。

通过上面的计算可以看出，如果非完全信息动态博弈稍微复杂，计算序列均衡就非常烦琐复杂，因而计算序列均衡除了技术之外，还需要一点灵感加运气。

第三节 经典举例

这一节将介绍一些经典例子，一方面表明博弈论的运用前景十分广阔，另一方面也有利于读者进一步掌握序列均衡的计算与内涵。

需要指出的是，序列均衡作为博弈论的核心均衡概念，在运用中体现出了巨大的理论威力，然而，它仍然存在着令人不满意的地方。首先，序列均衡的计算

十分的烦琐和困难,尤其当博弈存在信息不完全和具有动态特征时更是如此。更为要命的是,如何求解扩展式博弈的序列均衡至今仍没有一般通行的方法。常常需要先去把均衡猜出来,然后再检验猜测是否正确。这就是为什么博弈论发展到今天仍只能研究和处理一些非常简单的模型。

这一节分析的案例,基本上都满足完美贝叶斯均衡就是序列均衡这样的条件,目的在于简单。

一、信号博弈

在经济学中,信息不完全基本上又可等同于信息不对称,几乎所有的信息不完全模型都与信息不对称有关,逆向选择和道德风险就是信息不对称的两种表现形式。在这里我们将介绍信息不对称中的一个重要模型——信号博弈,信号博弈是研究逆向选择的一个重要模型,最早由斯彭思(Spence,1973)提出并加以讨论,斯彭思也因此获得了2001年的诺贝尔经济学奖。

在许多情况下,信息可能是"不对称的",例如,应聘者知道自己的真实能力但雇主不知道;卖二手车的人了解汽车的真实状况但买者不清楚;买保险的人可能知道自己的健康状况但保险公司不了解。这时,了解内情的人可能采取向未知者发送信号的方式来传递信息,而未知者会根据发信号者的行动来推断信号的真实性,以便各自谋取最大的利益。

在信号博弈中,通常将知情者称为信号发送者,而将未知者称为信号接收者,信号接收者只能观察到信号发送者的行动,而不能直接观察到信号发送者的类型。这里,我们将由易到难介绍几个著名的信号博弈模型。

(一)进入博弈

一个市场上存在一个在位企业和一个挑战企业,挑战企业可能实力强大,概率为p,也可能实力较弱,概率为$1-p$,满足$1>p>0$。挑战企业知道自己的类型(是否强大),但在位企业不知道。挑战企业可以选择积极备战,也可以选择不备战,而在位企业能够观察到挑战企业是否备战(但不能观察到挑战者的真实实力),并决定是否迎战和退缩。如果挑战者积极备战,那么比不备战需要花费更多的成本;如果挑战者强大,那么位者选择迎战将会损失惨重;如果挑战者弱小,那么在位者选择迎战将能够获得较大收益。所以当挑战者强大时,在位者会主动避免冲突,而挑战者也有动机向在位者传递自己强大的信息,但在位企业并不会轻易相信。为了分析方便,不妨设挑战企业为参与者1,在位企业为参与者2,自然决定挑战者为强或弱,分别用S和W表示,挑战企业的行动为备战和不备战,分别用R和U表示,在位企业的行动为迎战和退缩,分别用F和A来表

示。双方具体的收益细节如图 7-8 所示。实际上，挑战企业选择是否备战就是在向在位企业发送信号，而在位企业接收到这个信号后，会据此来选择相应的行动。该博弈的一般扩展式留作作业，由读者来完成。

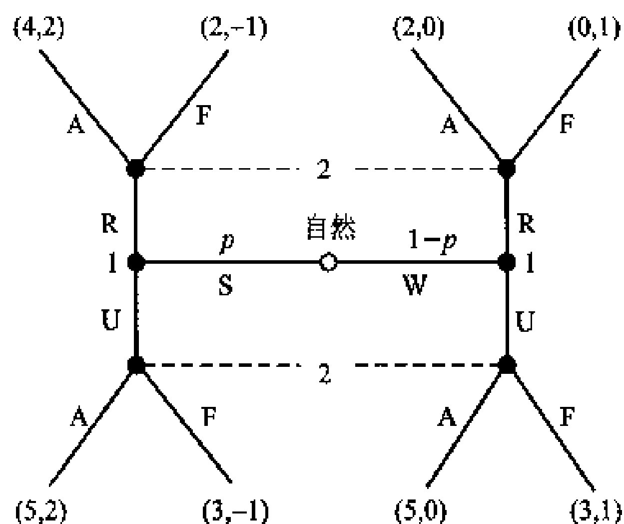


图 7-8　进入博弈

㊂ 习题 7.1　请给出上述进入博弈的一般扩展式。如若可能，请给出该博弈的一个序列均衡。

容易知道，挑战者有两个信息集{S}和{W}，因而存在 4 个"纯策略"，在位者也有两个信息集{SR，WR}和{SU，WU}，因而同样存在 4 个"纯策略"。那么该博弈的序列均衡是什么呢？经过观察，我们很容易发现，弱挑战者选择不备战是一个严格优策略，这就意味着一个弱挑战者将始终选择不备战。剩下的问题就只需要考虑挑战者为强时的情况，而强挑战者有两种行动可采取：

（1）强挑战者选择备战。这时，不妨假设在位者的策略为，如果观察到备战，那么选择 A；如果观察到不备战，那么选择 F。相一致的信念为，如果观察到备战，那么认为挑战者实力强大；如果观察到不备战，那么认为挑战者实力弱小。即，如果信息集{SR，WR}发生，那么在位者一定会认为发生的历史为 SR，而不会是 WR（因为弱挑战者选择不备战是严格优策略），因而其最优策略只能是 A。如果信息集{SU，WU}发生，那么在位者一定会认为发生的历史为 WU，而不会是 SU，因而其最优策略只能是 F。给定在位者的审时度势，如果强挑战者选择不备战，那么它的收益将从 4 变为 3，所以强挑战者不会偏离备战的选择。

（2）强挑战者选择不备战。这时情况有点复杂。由于挑战者无论强弱都将选择不备战，因而信息集{SR，WR}实际上不可能走到。对于在位者而言，其认为历史 SU 发生的概率为 p，历史 WU 发生的概率为 $1-p$。如果选择退缩，那么预期收益为 $2p$；如果选择迎战，那么预期收益为 $1-2p$。因此，当 $p\geqslant 1/4$ 时，最

优策略为退缩；当 $p \leq 1/4$ 时，最优策略为迎战。

当 $p \geq 1/4$ 时，在位者会选择退缩，这时强挑战者选择不备战是最优策略，因为选择备战收益为 $4 < 5$。由此可知，无论挑战者是强是弱都选择不备战，而在位者的信念为 SU 发生的概率为 p，WU 发生的概率为 $1 - p$，$p \geq 1/4$，策略为退缩。对于信息集{SR，WR}，由于其不可能发生，因而不访设 SR 发生的概率为 s，$1 \geq s \geq 0$，那么选择退缩的期望值为 $2s$，选择迎战的期望值为 $1 - 2s$。由此可得，当 $s \geq 1/4$ 时，最优策略为退缩；当 $s \leq 1/4$ 时，最优策略为迎战。

当 $p \leq 1/4$ 时，在位者会选择迎战，这时强挑战者选择不备战将获得 3 的收益，如果强挑战者选择备战（即发生背离），那么在在位者迎战的情况下收益为 2（其概率为大于等于 3/4），而在在位者退缩的情况下收益为 4（其概率为小于等于 1/4），所以其期望值小于等于 5/2，显然小于 3。因此，强挑战者仍会选择不备战。同理，对于信息集{SR，WR}，由于其不可能发生，信念可以任意设，只要策略与信念一致即可。

综上所述，进入博弈就存在两类序列均衡：

（1）挑战者的策略为（如果强，选择备战；如果弱，选择不备战），而在位者相信选择备战的挑战者实力强大，选择不备战的挑战者实力弱小，相应的策略就是（如果对手备战，那么选择退缩；如果对手不备战，那么选择迎战）。

（2）不论挑战者的类型是什么，挑战者都选择不备战，而在位者相信挑战者实力强大的概率为 p。如果 $p \geq 1/4$，那么在位者的策略就是退缩；如果 $p \leq 1/4$，那么在位者的策略就是迎战。在信息集{SR，WR}，如果 $\mu_2(SR) = s$，$1 \geq s \geq 0$，当 $s \geq 1/4$ 时，最优策略为退缩；当 $s \leq 1/4$ 时，最优策略为迎战。

在上而这个例于中，第 1 类均衡被称为**分离均衡**（separating equilibrium），所谓**分离均衡**是指不同类型的发信号者会发出不同的信号，因此信号接收者通过观察发信号者的行动就可以知道发信号者的类型。例如，强挑战者选择备战，弱挑战者选择不备战。而第 2 类均衡则被称为**混同均衡**（pooling equilibrium），所谓混同均衡是指所有类型的发信号者都发出相同的信号，因此信号接收者无法通过观察发信号者的行动来知晓发信号者的类型。例如，无论挑战企业强弱，它都发出一样的信号——不备战。介于分离均衡与混同均衡之间还存在各种各种的均衡，例如**半混同均衡**（semi - pooling），即有些类型能够区分，而有些类型无法区分，**杂合均衡**（hybrid），即某个类型的发信号者发出的信号为 m，而另一个类型的发信者则随机发出信号 m 和 m^*。

㊐ 习题 7.2　信号博弈中的分离与混同均衡。仍考虑进入博弈，但稍作修改，即改变全历史 WRA 和 WRF 下的收益值。当其为多少时，使得进入博弈只存在唯一一个序列均衡？且为分离均衡？当其为多少时，进入博弈只存在为唯

一一个序列均衡，且为混同均衡？

（二）体育赞助

一个市场上有些厂商生产的商品质量[1]好，而有些厂商生产的商品质量差，对于消费者而言，显然并不清楚哪个品牌的质量好而哪个品牌的质量差，对于厂商而言，如何向消费者发出信号，以表明自己生产的产品质量过硬呢？一个自然的想法就是厂商向体育比赛进行赞助，以表明自己的产品质量过硬，因为只有自己的产品过硬从而吸引回头客，进行体育赞助的花费才是值得的，如果产品质量不过硬即便进行体育赞助也无法赢得回头客。企业向消费者传递质量过硬的信号方式可以有多种形式，例如巨额广告、社会公益活动等。下面我们将构建一个动态贝叶斯博弈以揭示上述现象。

为了简单，我们不妨设市场上存在一个企业（F）和消费者（C），自然（Nature）决定企业生产出高质量的产品的概率为 q，低质量产品的概率为 $1-q$，H 表示高质量产品带给消费者高质的效用，L 表示低质量产品带给消费者低质的效用，企业知道自己产品的质量，但消费者不清楚，因而在第一阶段企业向消费者传递的信号为产品的价格 p_1 和体育赞助费 E，而消费者根据 (p_1, E) 来判断产品的质量，以决定是否购买，如果购买，那么消费者将彻底了解产品的质量状况，这时企业重新定价，而消费者再次决定是否购买该产品。为了简单，假设产品的数量始终为 1，高质量产品的成本为 c_H，低质量产品的成本为 c_L，$H > c_H > c_L > 0$，p_1 表示企业第一次要价，p_2^H 则表示企业的高质量产品的第二次要价，p_2^L 为低质量产品的第二次要价。很明显该博弈是一个两阶段博弈，需要注意的是消费者的策略，在第一阶段，消费者不清楚产品的质量，但如果其购买了商品，那么在第二阶段消费者将完全了解产品质量，因而消费者的策略较为复杂。同时还须注意，不能分割任何的泛子博弈。该博弈的博弈树如图 7－9 所示，图中的阴影三角形表示企业的定价 p_2 是连续的。

㊂ 习题 7.3　请写出该博弈的一般扩展式，以及消费者的一个策略＋信念。

显然该博弈存在很多序列均衡，但我们所关心的问题是：是否存在一个序列均衡，其结果为高质量的企业愿意支付体育赞助，而低质量的企业不愿意支付体育赞助，显然这是一个分离均衡。为了进一步简化，不妨设低质量的产品对消费者而言效用为 0，即 $L=0$。我们不妨考虑如下的审时度势组合：

---

① 这里的质量实际上具有非常广泛的含义，可能代表着企业的信誉、较高的社会责任感、较完善的客户服务等。

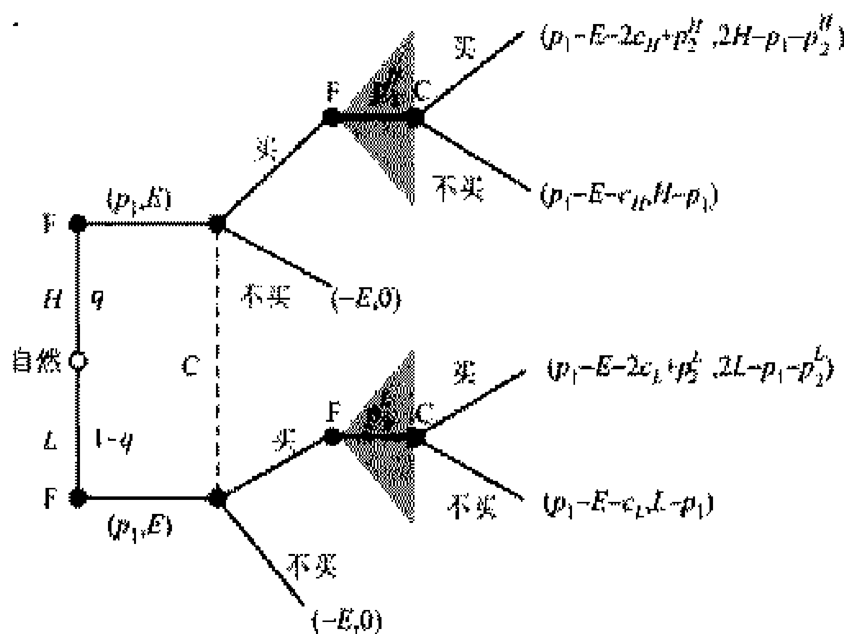


图 7－9　体育赞助博弈

企业的策略：如果第一期企业为高质量（即自然选择了 H），那么选择（p^{H*}，E^*）（即选择尽可能的高价和赞助体育），而第二期选择 $p_2^H=H$；如果第一期企业为低质量，那么选择（0，0）（即价格为0，不赞助体育），而在第二期选择 $p_2^L\geqslant 0$。

消费者的信念：消费者当且仅当在第一期观察到（$p^H\leqslant p^{H*}$，$E=E^*$）时才会相信企业生产的产品是高质量的，否则就认为企业生产的产品是低质量的。其中 $p^H\leqslant p^{H*}$ 表示尽管产品质量好但也不能漫天要价，否则低质量企业就会冒充高质量搞一锤子买卖，$E=E^*$ 表示体育赞助至少必须达到一定程度，如果 E 太小显然就不具备传递质量信号的意义。严格意义上来讲，应当是 $E\geqslant E^*$，但赞助费越小，企业的利润越高，所以取等号。

消费者的策略：在第一期，只有当消费者观察到（$p^H\leqslant p^{H*}$，$E=E^*$）时，消费者才会购买产品，否则不会购买；在第二期，如果价格不高于 H，那么购买，否则不会购买。如果第一期产品质量为 L，万一走到了第二期，那么只有当产品的价格为0，才会购买，否则不会购买。

首先，运用逆推法。可知，对于子博弈 $\Gamma(p_2^H)$，消费者选择买要好过不买，必须满足 $H\geqslant p_2^H$，因为 p_2^H 是连续的，所以不妨取 $H=p_2^H$，即只有价格不高于 H 时，消费者才会购买（价格越低对消费者越好，但企业越差）。对于子博弈 $\Gamma(p_2^L)$，消费者选择买要好过不买，必须满足 $L\geqslant p_2^L$，因为 p_2^L 不可能为负数，所以只有当 $p_2^L=0$ 时，消费者才会购买。这也说明了我们为什么会那样设置企业与消费者的策略。

接下来将依次分析企业与消费者的上述策略，在与信念保持一致的情况下，成为序列均衡的条件。

如果企业的类型为 H，那么企业给出(p^{H*},E^*)，其利润为 $p^{H*}+H-E-2c_H$。如果背离，那么有两种可能：一种是在第二阶段要价高于 H，利润为 $p^{H*}-E-c_H$；一种是在第一阶段给出的(p_1,E)不等于(p^H,E^*)，利润为0。因而，如果该策略为最优策略就必须满足

$$p^{H*}+H-E^*-2c_H\geqslant 0。$$

如果企业的类型为 L，那么企业给出(0,0)，其利润为0。如果背离，那么有几种情况：一种是 $E<E^*$，一种是 $p_1>p^{H*}$，$E=E^*$。无论哪种情况，都意味着消费者认为产品为低质量，不会购买，利润不可能大于0；另一种情况，企业冒充自己是高质量企业，那么它必然选择$(p_1\leqslant p^{H*},E=E^*)$，如果消费者购买了，那么将发现产品为低质量，因而第二期不可能在购买，所以企业的利润不超过 $p^{H*}-E^*$，因而，如果该策略为最优策略，必须满足

$$0\geqslant p^{H*}-E^*\rightarrow E^*\geqslant p^{H*}$$

对于消费者而言，假如观察到$(p\leqslant p^{H*},E=E^*)$，那么他相信厂商是高质量厂商，购买商品将使其得到的收益不小于 $H-p^{H*}$（第二阶段的收益为0，$p_2^H=H$，消费者购买与不购买收益都一样，所以我们不妨假设消费者仍然购买），而如果其不购买，那么收益为0。因而，在消费者观察到$(p\leqslant p^{H*},E=E^*)$时，如果该策略是最优策略，那么必须满足

$$H-p^{H*}\geqslant 0\rightarrow H\geqslant p^{H*}$$

如果消费者观察到其他的价格与赞助费，他就会认为厂商是低质量厂商，不购买该商品，收益为0，相反，如果购买，那么收益为 $0-p_1$，因而最优反应就是不购买该商品。如果产品为低质量，万一走到了第二期，那么消费者只有当 $p_2^L=0$ 时，才可能购买。

综上所述，上述审时度势为一个序列均衡，当且仅当

$$\min\{E^*,H\}\geqslant p^{H*}\geqslant E^*+2c_H-H$$

显然，如果 $H>2c_H$，$E^*>0$，那么就存在组合(p^{H*},E^*)满足上述条件。因而，一个高质量企业将会赞助体育，因为如果它不这样做，消费者就会认为其为低质量厂商，不会在第一期购买其产品，因而也就失去了解该厂商实际上是一个高质量厂商的机会，从而在第二期也不会购买其产品。为了防止低质量厂商冒充高质量厂商，E^* 的值应当足够高，从而使得低质量厂商这样做将出现负的利润。

体育赞助博弈之所以能够得出上述结论，依赖于该博弈的两个重要结构：一

是存在回头客,如果没有回头客,那么低质量厂商始终有动机去冒充高质量厂商,由于 $c_H > c_L$,因而通过模仿高质量厂商,低质量厂商始终能够获得不低于高质量厂商的利润,因而在一个很少产生回头客的行业,各种假冒伪劣产品更容易泛滥,而回头客较高的行业,假冒伪劣产品就会少得多。例如,在北京的风景名胜区,由于消费者大多都是旅客(所以不是回头客),因而充斥着各种低劣产品和服务,而不得不通过政府的力量来加以监管;而在北京的大胡同,由于消费者通常为周围的居民(熟客),所以低劣产品就远远少得多。二是消费者的选择是相当极端的(即只有观察到(p^{H^*},E^*)才买,否则不买),有可能即使消费者观察到(p_1,$E \neq E^*$)仍然会购买,这时,高质量厂商可能不愿支出赞助费,因为其不进行体育赞助仍有产品被消费者所购买的希望,从而消费者发现产品为高质量,因而也能产生回头客。这就解释了为什么仍有很多的低质量厂商存在,因为存在混同均衡的可能。

㊐ 习题 7.4　请你证明体育赞助博弈还存在着混同均衡,即博弈的第一阶段高质量厂商和低质量厂商都选择一样的价格和赞助费。

(三) 文凭的信息价值

近几年中国掀起了上研究生的热潮,大多数学子上研究生的目的在于能够获得一个高薪职务,较少有人是为了做学问而选择继续深造。人们为什么会把研究生文凭看作是获得高薪的敲门砖呢? 一个想法是,由于人们的能力除了自己清楚以外,其他人是不清楚的,而文凭可能包含着能力强这样的信息,因为能力强的人获得研究生文凭较能力弱的人所费成本要低,因而通过获取较高的文凭来向雇主传递能力强这样的信号,由于能力强的人具有较高的生产率,因而应当获得较高工资。为了简单,我们假设一位劳动者其能力有可能高(H),也有可能低(L),$H > L$,能力的高低由自然决定,$\rho(H) = p, 1 \geq p \geq 0$。能力越高,获得教育程度 e 所付出的代价越低,不妨用 e/K 表示,K 代表劳动者的能力,$K = H$ 或 L。劳动力市场上存在 n 个企业,定义为 $1,2,\cdots,n$。企业对劳动者的能力并不清楚,而只能观察到劳动者的教育程度,企业付给劳动者的工资用 w 表示,因而劳动者的收益为 $w - e/K$,企业的收益为 $K - w$。由于市场上企业之间的竞争较为充分,因而企业的收益 $K - w = 0$,因为如果有企业的收益 $K - w > 0$,那么就有企业会开出更高的工资,以便将该劳动者吸引到该企业来,这也是我们在模型中引入多个企业的原因。该动态博弈的博弈树如图 7 - 10 所示,图中的阴影三角形表示($w_1,\cdots,w_n$)为连续向量。由于存在多个企业,工资有多个,我们用 w_1, w_i, w_n 来简单地表示劳动者从工资向量($w_1,\cdots,w_n$)中选择一个工资。

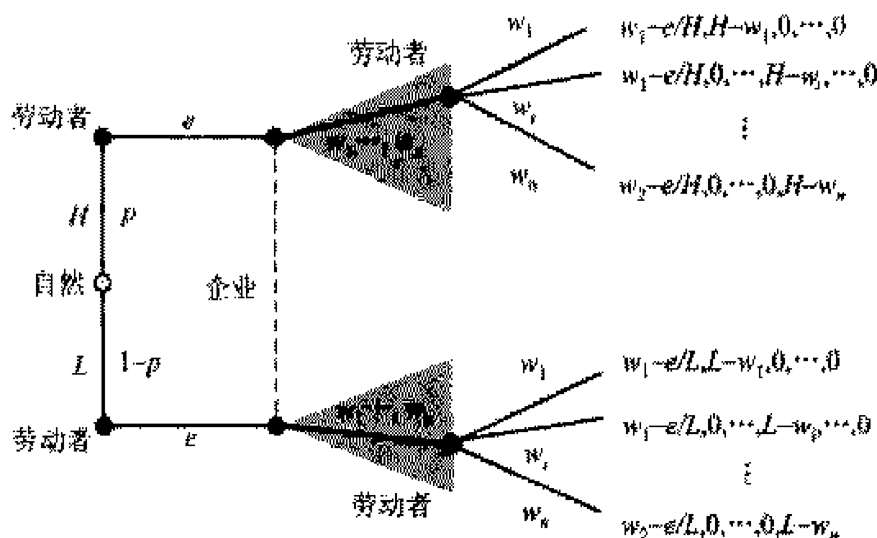


图 7-10　教育博弈

(易) 习题 7.5　请给出教育博弈的一般扩展式，如果可能请给出一个序列均衡。

接下来要做的，就是证明在教育博弈这种模型中，高能力的人将选择一个较高的教育程度，以便和能力差的人相区别，这是一个序列均衡结果。明确目标之后，就容易分析了。不妨考虑如下审时度势组合：

劳动者的策略：如果是高能力，那么选择 $e=e^*>0$；如果是低能力，那么选择$e=0$。在第二阶段，则选择 $\max\{w_1,\cdots,w_n\}$，即最高的工资；如果最高工资多于1个，则在最高工资中随机选择。

企业的信念：如果观察到劳动者选择教育程度 e^*，那么认为劳动者是高能力；如果观察到劳动者选择教育程度 $e\neq e^*$，那么认为劳动者是低能力。

企业的策略：如果认为劳动者为高能力，那么给出高工资 $w=H$；如果认为劳动者为低能力，那么给出低工资 $w=L$。

首先，我们运用逆推法，可知由于企业之间相互竞争（我们假设完全竞争），因而各企业实际上给出的工资都是一样的，即 $w=K$，这实际上就是所谓的均衡工资，企业利润为 0。因为如果企业所给工资低于均衡工资，那么它将招不到工人；如果所给工资高于均衡工资，那么它将吸引到所有工人，因而降低工资是有利可图的。

对劳动者而言，给定企业的审时度势，在博弈的第二阶段，无论其能力高低，选择最高工资始终是最优反应。在第一阶段，如果高能力者选择 e^*，而不是选择 0，那么必须满足

$$H-e^*/H\geqslant L$$

否则高能力者就有可能假装成低能力者。同时，如果低能力者选择 0，而不是选择 e^*，那么必须满足

$$L - 0 \geq H - e^*/L$$

否则低能力者就有可能假装成高能力者。

给定劳动者的策略，那么企业如果观察到 $e = e^*$，那么给出高工资，因为给出低工资并不能增加企业的利润，因为其他企业会给出高工资，从而使得企业不可能雇用到劳动者。另一方面，如果观察到 $e \neq e^*$，那么给出低工资，因为其他所有企业会给出低工资，它没有必要给出高工资，第二就是（有可能）劳动者是一个低能力者，给出高工资将意味着利润是负的，即 $L - H < 0$。所以劳动者的上述策略 + 信念是企业策略的最优反应。

综上所述，上述审时度势组合为一个序列均衡，当且仅当

$$H(H - L) \geq e^* \geq L(H - L)$$

因为 $H > L$，所以这样的 e^* 是存在的。如果高能力获得的工资越高，低能力获得的工资越低，即 $H - L$ 的值越大，那么 e^* 的值也越大，表示受教育的程度越高。近 10 年来，拥有硕士学位者与本科学位者的收入差距平均而言是在不断扩大的，这就从一个侧面揭示了为什么近 10 年来中国会发生研究生热。近年，由于研究生扩招，使得拥有硕士学位者的平均收入下滑较大，从而出现了 2007 年报考研究生人数的大幅下降。

上述模型虽然简单，但并不等于其分析力就会减弱，如果我们引入政府对高能力者征收重税（因为其收入较高），那么有可能高能力者就会将自己假装成低能力者，所以在税收制度的设计上，同样面临效率与公平难以兼得这样的困难。经济学家莫里斯（Morlees）获得 1996 年诺贝尔经济学奖的一个重要贡献，就是其运用博弈论的方法严格证明了既能实现效率、又能实现公平这样的最优税制实际上是不存在的。

㊥ 习题 7.6　在上述教育博弈中，如果引入政府对高能力者征收重税，比如规定当工资达到一定程度 $w \geq H^*$，那么将征收 tw 的税收，$1 \geq t \geq 0$，那么高能力者将选择 $e = 0$ 成为一个序列均衡结果的条件是什么？（提示：证明如果不满足分离均衡，那么就意味着满足混同均衡。）

二、阻吓定价

在初级微观经济学中，垄断厂商的定价是以获得最大垄断利润为原则，其一个重要条件就是垄断厂商能够阻止新厂商进入该行业。如果做不到，新厂商就会加入该行业，使得垄断利润消失。形成进入障碍的可能性有专利、政府特许，但也有许多是源自垄断厂商的自然力量。这种自然力量通常有两种：一是实行所谓的限制性定价（limit pricing）。垄断厂商通过制定一个较低的价格，让潜在的进入者认为该行业无利可图，从而打消进入该行业的企图。限制性定价通常

在两种情况下产生:进入该行业需要巨额的固定投资,例如汽车、航空、石油等行业;或是潜在进入者对该行业内某个影响利润的因素(例如生产的平均成本或市场需求)具有不确定性。第二种做法是,潜在进入者虽然知道垄断行业存在超额利润,但却迫于在位厂商的威胁,不敢贸然进入。因为在位厂商威胁一旦有新进入者,它将把价格压至最低,进行价格战,让进入者(和自己)遭受巨大损失,这种策略就是所谓的阻吓定价。

考虑如图 7 - 11 所示的一个非常简单的进入模型。

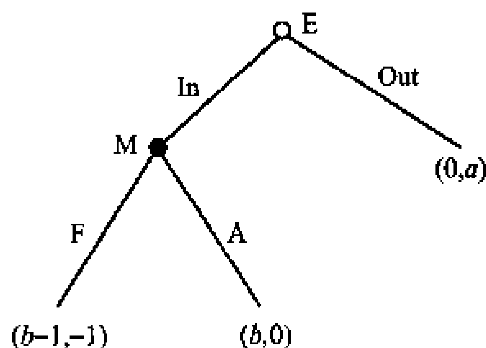


图 7 - 11 简单进入模型

新厂商(E)决定要不要进入垄断市场。若它选择不进入(Out),则市场内原来的垄断者(M)得到垄断利润 $a>1$,而 E 的利润为 0。若 E 选择进入垄断行业(In),则垄断者将面临是否进行价格战的选择。如果 M 选择和平共处(A),则由于垄断利润被瓜分,所以假设 M 的垄断利润降为 0,而 E 能获得 b 的垄断利润,其中 $1>b>0$。如果 M 选择进行价格战(F),则利润为 -1,而 E 得到的利润为 $b-1$。可知该博弈唯一的纳什均衡是新厂商选择进入(In),而垄断厂商选择和平共处(A)。如果进入博弈重复进行 T 期(不是无限期),那么根据我们在前面分析重复博弈时得出的结论,在垄断厂商的威胁是不可置信的情况下,即使博弈进行多期,唯一的子博弈完美均衡仍然是新厂商每一期都进入,而垄断厂商选择与新进入者和平共处,即阶段博弈的纳什均衡在动态博弈中重复 T 次。然而现实与上述理论分析却不一致,我们通常观察到的现象是垄断厂商往往能够有效阻吓新厂商的进入。一个解决的办法是引入无限期,由于博弈重复进行无限期,因而垄断厂商为了长期利润就会不惜代价发动价格战,从而有效阻吓新厂商的进入。但现实中仍有大量事实表明,即使博弈重复有限期,垄断厂商在新厂商进入垄断行业后仍会发动价格战,从而有效阻吓进入者,这种理论与现实的矛盾就是所谓的**连锁店悖论**(chain-store paradox)。

针对连锁店悖论,克瑞普斯和威尔逊提出的思路是[①],由于存在着信息不对称,垄断厂商发动价格战,从而使进入者认为自己"疯狂",因而阻止新厂商的进入。他们的模型假设在一个垄断市场,垄断者可区分为两种类型:其一是理性厂商,称作正常型(normal type)厂商;另一种是喜好不惜成本打价格战,而不喜欢

① Kreps D and Wilson R. "Reputation and Imperfect Information". *Journal of Economic Theory*. 1982,27:253 ~ 279

和平共处的厂商,称作疯狂型(crazy type)厂商。假设垄断厂商为疯狂型厂商的概率为δ,而正常型厂商的概率为$1-\delta$。那么整个博弈的扩展式如图7-12所示。

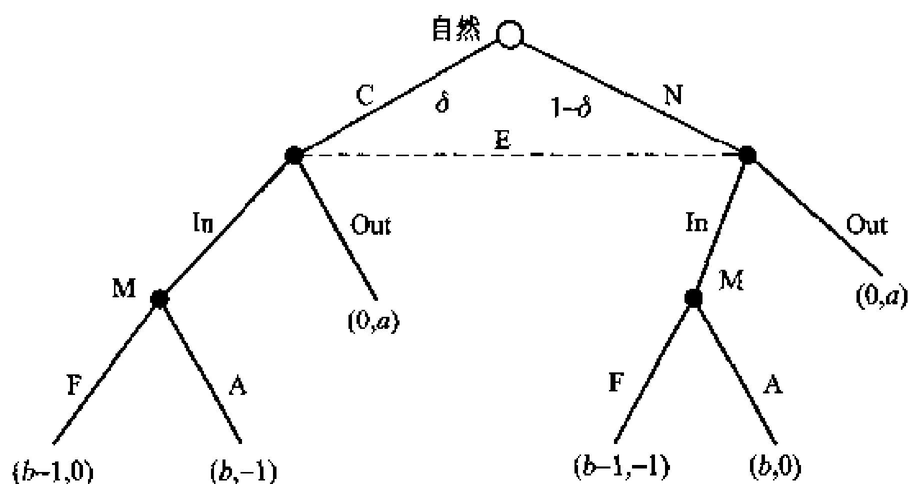


图7-12　进入模型

其中C表示疯狂,N表示正常。图7-12的左半部分和右半部分唯一不同之处,在于垄断者选择F(战斗)和A(和平共处)的收益正好相反。很明显,疯狂型垄断厂商在遇到新进入者时,一定会进行价格战。在这个设定下,克瑞普斯和威尔逊对于价格战的解释如下:在进入博弈重复多期的情况下①,由于进入者不知道垄断者的类型,因此正常型垄断者有伪装成疯狂型厂商的动机。伪装的方法,就是在新厂商进入时,即进行价格战。虽然这会受到损失,但如果因此使后来者误认为它是疯狂型而放弃进入,那么牺牲短期利润进行价格战以换取长期利润就是值得的。因此如果垄断者在市场上存在足够多期,那么在最初的几期里,任何类型的垄断厂商在遇到新进入者时,一定会发动价格战,因而没有厂商愿意进入。只有到最后几期,由于正常型厂商为了伪装所付出的代价已不足以用长期利润(剩下的几期利润)来弥补时,才会暴露自己的真实类型。进入者也才会在最后几期真正知道垄断者的类型。只有它发现垄断者是正常型厂商时,才会选择进入。

假设有一个N期的垄断市场,垄断者M在这段期间内会留在市场中。M的收益为每一期的收益的总和。现在有N个潜在的进入者想进入垄断市场,以$1,\cdots,n,\cdots,N-1,N$表示。也就是说,在第n期,厂商n和垄断厂商进行一次如图7-12所示的博弈。第n期,只有厂商n有进入垄断市场的选择。厂商n的

① 还记得我们在介绍重复博弈时说过,阶段博弈不仅仅是静态博弈,它本身可以是非完全信息动态博弈。

收益即是它在当期进行该博弈的收益。令 p_n 为厂商 n 在进人前，认为 M 属于疯狂型厂商的概率，p_n 的值会随每期垄断者的行为而改变。因此，厂商 n 的策略，是根据它的猜测来决定要不要进人该市场，而 M 的策略，是在每一期，如果有厂商进人，它根据 p_n 的值来决定要不要进行价格战。为了回答连锁店悖论，我们不妨考虑如下的审时度势组合：

垄断者的策略：

M1 假如垄断者属于疯狂型，则只要有新厂商进人该市场就进行价格战。

M2 假如垄断者属于正常型，则在最后一期（$n=N$），垄断者会选择和新进人者和平共处。当 $n<N$ 时，如 $p_n>b^{N-n}$，那么它一定进行价格战；如 $p_n\leqslant b^{N-n}$，那么它进行价格战的概率是 $(1-b^{N-n})p_n/(1-p_n)b^{N-n}$，即如果 $p_n\leqslant b^{N-n}$，那么垄断者使用混合策略 $\{(1-b^{N-n})p_n/(1-p_n)b^{N-n},(b^{N-n}-p_n)/(1-p_n)b^{N-n}\}$。

新厂商的信念：

B1 $p_1=\delta$。

B2 在任何 $n<N$ 的时期中，如厂商 $n-1$ 进人市场而垄断厂商采用和平共处策略 A，或 $p_{n-1}=0$，则 $p_n=0$。

B3 $n>1$，如 $p_{n-1}\neq 0$，而在第 $n-1$ 期时，新厂商 $n-1$ 进人市场而垄断者进行价格战，则 $p_n=\max(b^{N-n+1},p_{n-1})$。

新厂商 n 的策略：

E1 如果 $p_n>b^{N-n+1}$，则不进人市场（Out）。

E2 如果 $p_n<b^{N-n+1}$，则进人市场（In）。

E3 如果 $p_n=b^{N-n+1}$，则使用混合策略（$1/a$Out，$(1-1/a)$In）。

如果证明上述审时度势组合为序列均衡，那就意味着在理论上为连锁店悖论提供了一种合理的解释。

证明：B1 不证自明，因为它是先验概率。B2 也很明显，因为如果垄断厂商选择和平共处，那么它一定是正常型厂商，所以 $p_n=0$。B3 的证明如下：如果 $p_{n-1}>b^{N-n+1}$，则根据垄断者的策略 M2，它应该在新厂商 $n-1$ 进人时，进行价格战。因此如果垄断厂商发动价格战，那么厂商 n 没有获得额外的信息，不会改变它对垄断者的猜测，因而 $p_n=p_{n-1}$。如果 $p_{n-1}\leqslant b^{N-n+1}$，那么根据 M2，垄断者应该在厂商 $n-1$ 进人时，用 $(1-b^{N-n+1})p_{n-1}/(1-p_{n-1})b^{N-n+1}$ 的概率进行价格战。因此，利用贝叶斯法则可得

p_n =（M 是疯狂型的概率）/（M 打价格战的概率）

=Prob（M 是疯狂型）×Prob（M 进行价格战|M 是疯狂型）÷{Prob（M 是疯狂型）×Prob（M 进行价格战|M 是疯狂型）+Prob（M 是正常型）×

Prob(M 进行价格战|M 是正常型)}

$$=\frac{p_{n-1}\times 1}{p_{n-1}\times 1+(1-p_{n-1})\dfrac{(1-b^{N-n+1})p_{n-1}}{(1-p_{n-1})b^{N-n+1}}}=b^{N-n+1}$$

因此 p_n 总是等于 p_{n-1} 和 b^{N-n+1} 中较大的一个。

E1、E2、E3 的证明如下：

(1) 若 $p_n>b^{N-n}$，则由 M2 可知，新厂商 n 进入后，无论是什么类型的垄断者都将进行价格战，因此新厂商选择不进入。

(2) 若 $p_n\leqslant b^{N-n}$，则由于新厂商 n 进入后，正常型垄断者会用 $(1-b^{N-n})p_n/(1-p_n)b^{N-n}$ 的概率进行价格战，因此新厂商进入市场的预期收益为

$$p_n(b-1)+(1-p_n)\left[\frac{(1-b^{N-n})p_n}{(1-p_n)b^{N-n}}(b-1)+\frac{(b^{N-n}-p_n)}{(1-p_n)b^{N-n}}b\right]$$
$$=(b^{N-n+1}-p_n)/b^{N-n}$$

如果 $p_n>b^{N-n+1}$，那么新厂商同样选择不进入。由于 $b^{N-n+1}<b^{N-n}$，综合(1)、(2)可知，当 $p_n>b^{N-n+1}$ 时，新厂商肯定不进入；当 $p_n<b^{N-n+1}$ 时，新厂商进入；当 $p_n=b^{N-n+1}$，则进入与不进入的预期收益相同，因此新厂商的策略为混合策略。为了计算的方便，不妨设其以 $1/a$ 的概率选择不进入。

最后剩下的，就是证明 M1 和 M2 是给定新厂商上述审时度势的最优反应。M1 显然成立，我们只需证明 M2(即垄断厂商为正常厂商)为最优反应。

(1) 若 $n=N$，即最后一期，正常型厂商和新厂商和平共处是最优策略(已经没有战斗的必要)，收益为 0。

(2) 若 $n<N$，则令 $V_{up}(n)$ 是当 $p_n>b^{N-n+1}$ 时正常厂商在第 n 期的预期利润，$V_{down}(n)$ 是 $p_n<b^{N-n+1}$ 时正常厂商在第 n 期的预期利润，$V(n)$ 则是 $p_n=b^{N-n+1}$ 时正常厂商的预期利润。很明显，如果 $p_n>b^{N-n+1}$，那么新厂商不会进入(E1)，因此垄断者可以一直获得 a 的利润，而 p_n 的值会维持不变(因为没有厂商进入，因而后来的新厂商无法通过观察以前的历史来修正对垄断厂商类型的猜测)。但 b^{N-n+1} 会随着时间的推移(n 变大)而增大。令 $m^*=\min\{m|p_1=\delta\leqslant b^{N-m}\}$，即 m^* 为大于先验概率的临界期，那么对第 m^* 期之前的任一期 $n<m^*$，由于 $p_n=p_1>b^{N-n}$，所以不会有新厂商进入(因为所有类型的垄断者都选择价格战)，垄断厂商会一直独享利润 a。在第 m^* 期，由于 $b^{N-m^*+1}<p_1\leqslant b^{N-m^*}$，垄断厂商要么战斗，要么采用混合策略，而新厂商不会进入，所以垄断者仍能获得 a 的利润。到了第 m^*+1 期，新厂商 m^*+1 开始有可能进入。假如 $p_{m^*+1}=b^{N-m^*}$，垄断者的利润为 $V(m^*+1)$。若 $p_{m^*+1}<b^{N-m^*}$，那么垄断者的利润为 $V_{down}(m^*+1)$。

因此

$$V_{up}(n)=(m^*-n)a+a+V_{down}(m^*+1) \quad 其中 n<m^* \tag{7-B-1}$$

$$V(n)=(1/a)[a+V_{down}(n+1)]+(1-1/a)V_{down}(n+1) \\ =1+V_{down}(n+1) \tag{7-B-2}$$

$$V_{down}(n)=\begin{cases}-1+V(n+1) & 若 M 进行价格战 \\ 0 & 若 M 选择和平共处\end{cases} \tag{7-B-3}$$

我们先看式(7-B-3)。根据B3,可知当垄断厂商进行价格战时,它在当期损失1,但会使后面的新进入厂商修正猜测,即 $p_{n+1}=\max(b^{N-n},p_n)$,而当 $p_n<b^{N-n+1}$ 时新厂商才可能在第 n 期进入,所以 $p_{n+1}=b^{N-n}>b^{N-n+1}$;当垄断厂商选择和平共处时,收益为0,所以得到(7-B-3)式。

再看式(7-B-2)。当第 n 厂商采取混合策略($1/a$Out,$(1-a)$In)时,如若不进入,垄断厂商在 n 期的收益为 a,同时后面的新进入看对垄断厂商类型的猜测不会改变,即 $p_{n+1}=p_1=\delta=b^{N-n+1}$,但另一方面 b^{N-n+1} 又会随 n 的增加面增大,即在 $n+1$ 期,$p_{n+1}<b^{N-n}$,即引起第 $n+1$ 厂商进入,所以垄断厂商在 n 期以后的收益为 $V_{down}(n+1)$;如若进入,那么垄断企业将采用混合策略($p_n=b^{N-n+1}<b^{N-n}$),在 n 期收益等于0,如果垄断厂商在 n 期发动价格战,第 $n+1$ 厂商的猜测变为 $p_{n+1}=b^{N-n}$,如果垄断厂商采取和平共处,那么 $p_{n+1}=0<b^{N-n}$,两者都意味着会引起第 $n+1$ 厂商的进入,所以垄断厂商在 n 期以后的收益为 $V_{down}(n+1)$。综上所述,即得到(7-B-2)式。

利用迭代法,以及最后一期和平共处的结论,很容易解出 $V_{down}(N)=0$,从面 $V_{up}(n)=(m^*-n+1)a$,其中 $n\leqslant m^*$;$V(n)=1$;$V_{down}(n)=0$。因此当一个垄断者在新厂商 n 进入时,如果它进行价格战,那么由于 $p_{n+1}=\max(b^{N-n},p_n)$,所以它的利润为

$$-1+V(n+1)=0 \quad 若 p_n\leqslant b^{N-n}$$

$$-1+V_{up}(n+1)=(m^*-n)a-1, m^*>n, a>1 \quad 若 p_n>b^{N-n}$$①

而如果它不进行价格战,利润为0。由此可知,当 $p_n\leqslant b^{N-n}$,是否进行价格战对垄断厂商而言都一样(等于0),所以它用混合策略。而当 $p_n>b^{N-n}$ 时,进行价格战的预期利润大于0,因此它一定会进行价格战。这就证明了M2是垄断厂商的最优策略。

综上所述,上述垄断厂商和进入厂商的审时度势组合是阻吓定价博弈的

---

① 记住 $p_{n+1}=\max(b^{N-n},p_n>b^{N-n})$

一个序列均衡。如果进入者相信垄断者可能会是不理性的(即使这个概率很小),那么垄断者利用这个想法进行价格战以阻吓新厂商进入的做法,就是相当合理的。同时,如果期数很多,那么 b^N 会趋于0,从而小于 $p_1=\delta$,因此,最初的几期新厂商都不会进入垄断市场。此外,当 $p_1=\delta$ 的值越大,m^* 也越大,垄断厂商获得的垄断利润也越多。这就从理论上为连锁店悖论提供了一种合理的解释。

(难) 习题7.7 (信息不对称有限重复博弈中合作的再讨论)实际上,阻吓定价博弈的思路可以用来解释有限重复信用困境中同样也能出现合作解,因为由于信息不对称,商人1对商人2是有仇必报的商人还是理性商人摸不准,这就会使得(在不附加任何条件的情况下)合作解在最初的几期中出现。请你试图分析上述想法是否正确?

三、诉讼

诉讼博弈包含了博弈论研究中的一些重要方面,例如信息不对称、信息揭露、谈判等,都可以在诉讼模型中自然地出现。诉讼中的原告和被告通常都拥有不同的私人信息,而法律中对原、被告双方的举证责任有相当严格和细致的规定。在诉讼过程中,双方提供的证据实际上就是一个信息揭露的过程,同时也是一个讨价还价的谈判过程,因而诉讼模型具有相当的代表性。这里我们考虑一个相当简化,但已包含上述所有特征的诉讼模型(Nalebuff,1987)。

假设发生了一次意外事故,受害者决定对肇事者提出赔偿要求。被告所应承当的赔偿由 w 和 q 决定,其中 w 是被告对原告所造成的损失,且原告和被告双方都知道 w 的值,q 则是被告应负担责任的比例,我们假设 q 的值是被告的私人信息。因此被告在这件事中,所应承当的赔偿金额为 qw。诉讼的程序如下:原告先提出一个赔偿金额 s,如果被告接受,即达成庭外和解,博弈结束。如果被告拒绝这个提议,那么原告可以选择是否诉讼。假如原告认为自己胜诉的机会很低,那么他可以放弃诉讼,各自的收益为0。如果原告决定诉讼,那么假设 q 的值会在法庭上被揭露出来,这时被告必须赔偿原告 qw 的金额,同时他们各自承担诉讼成本,所以原告实际上的收益为 $qw-C_q$,而被告的收益为 $-qw-C_d$,其中 C_q 和 C_d 是原告和被告各自的诉讼成本,都大于0。这个诉讼博弈可以用图7-13的博弈树来表示[①]。其中P代表事件的原告,D代表事件的被告

① 这个模型出自 Nalebuff,B. 1987. "Credible Pretrial Negotiation." *Rand Journal of Economics* 18:198-210

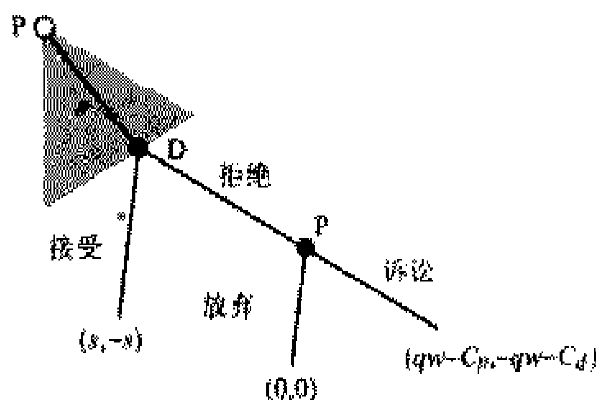


图 7－13　诉讼博弈

为了简化计算，我们假设：

A1　q 的值是在$[0,\bar{q}]$之间的均匀分布，它的密度函数为$f(q)$。如果$\bar{q}=1$，那么 q 就是需要承当责任的百分比。

A2　对原告来说，这个官司是值得打的，即 $E(q)w>C_p$，即获得赔偿的期望值大于诉讼成本，实际上这是一个先验态度。

诉讼博弈的关键是需要弄清楚原告有关 $E(q\mid s$ 被拒绝$)$的信念。

首先看信息集$\{s$ 拒绝$\}$以下的子博弈。给定原告的提议 s 已经被被告拒绝，那么原告就需要决定是否值得打官司，而这取决于他打官司的预期收益 $E(q|s$被拒绝$)w$ 和成本 C_p 的相对大小。若 $E(q\mid s$ 被拒绝$)w>C_p$，则官司值得打；若 $E(q|s$ 被拒绝$)w<C_p$，则不值得打这个官司，因此原告的最优策略为

如果 $E(q|s$ 被拒绝$)w>C_p$，则打官司；

如果 $E(q|s$ 被拒绝$)w=C_p$，随机选择打官司或放弃；

如果 $E(q|s$ 被拒绝$)w<C_p$，则放弃。

$E(q|s$ 被拒绝$)$实际上是原告关于 q 的信念，它的形成直接受到被告行为的影响，例如，如果原告提议的 s 很小，而被告加以拒绝，那么 q 非常小的概率就很大（即很可能错不在被告），因为被告要冒上法庭的风险而去拒绝一个较低的赔偿要求，如若上法庭，被告至少要支付 C_d 的诉讼成本。这说明 $E(q|s$ 被拒绝$)$是原告提议的 s 的函数，因此 $E(q|s$ 被拒绝$)$的值必须从被告在信息集$\{s\}$的选择来推算。

其次，被告是否接受提议 s，主要取决于 s 与 $r(s)(qw+C_d)$的相对大小，其中 $r(s)$为一旦被告拒绝 s 后，原告打官司的概率。如果 s 非常小，那么被告应接受和解要求；如果 s 非常大，那么被告会拒绝。令 $q(s)=(s/r(s)-C_d)/w$，即$q(s)$是刚好令 $s=r(s)(qw+C_d)$的 q 值，因此它是被告是否接受提议 s 的责任临界值。如果 $q>q(s)$，被告的最优策略是接受 s；如果 $q<q(s)$，则应拒绝。因此我们可以用贝叶斯法则来计算，当 s 被拒绝时，原告对被告责任值 q 的事后猜测为

$$E(q\mid s\text{被拒绝}) = \frac{\int_0^{q(s)} qf(q)\mathrm{d}q}{F(q(s))} \qquad (7-\mathrm{C}-1)$$

到此我们就得出与原告在信息集{s 拒绝}的策略相一致的信念 $E(q|s$ 被拒绝)。原告在被告拒绝后的选择,就变成计算 $\int_0^{q(s)} \frac{qf(q)}{F(q(s))}\mathrm{d}q \times w - C_p$ 是否大于0。如果大于0,则预期收益 $\int_0^{q(s)} \frac{qf(q)}{F(q(s))}\mathrm{d}q \times w$ 大于诉讼成本 C_p,打官司是值得的,否则放弃。由于 $\int_0^{q(s)} \frac{qf(q)}{F(q(s))}\mathrm{d}q \times w$ 是 $q(s)$ 的递增函数,$q(s)$ 越大,预期收益也就越大,因此一定存在一个 q^*,使得原告打官司与不打官司的收益相同,即 q^* 是下式的解

$$\int_0^{q(s)} \frac{qf(q)}{F(q(s))}\mathrm{d}q \times w = C_p \qquad (7-\mathrm{C}-2)$$

若 $q(s)=q^*$,则原告打官司与放弃的收益相同;若 $q(s)>q^*$,则会告下去;若 $q(s)<q^*$,则放弃,因而在信息集{s 拒绝}上,原告的最优策略如下

$$\begin{aligned} &\text{若 } q(s) > q^*, \quad \text{则 } r(s) = 1 \\ &\text{若 } q(s) < q^*, \quad \text{则 } r(s) = 0 \qquad (7-\mathrm{C}-3) \\ &\text{若 } q(s) = q^*, \quad \text{则 } r(s) \in [0,1] \end{aligned}$$

相应的信念由式(7-C-1)表示。

到此我们不妨猜测:

猜测1 当 $s>C_d$ 时,诉讼博弈进行到信息集{s 拒绝}的唯一均衡是:

被告的策略:若 $s>C_d+q^*w$,则所有 q 值大于或等于 $(s-C_d)/w$ 的被告会接受 s,而所有 q 值小于 $(s-C_d)/w$ 则会拒绝;若 $s\leqslant C_d+q^*w$,则 q 值大于或等于 q^* 的被告会接受 s^*,而所有 q 值小于 q^* 的被告则会拒绝。

原告的策略:开始提议 s,如果 s 被拒绝,若 $s>C_d+q^*w$,则 $r(s)=1$,若 $s\leqslant C_d+q^*w$,则 $r(s)=s/(C_d+q^*w)$。

原告的信念:由式(7-C-1)表示。

猜测1表明,如果和解提议 $s>C_d+q^*w$,那么只有责任值 $q>(s^*-C_d)/w$ 的被告会接受,否则拒绝,一旦 s 被拒绝,原告一定会告下去。如果和解提议 s^* 小于或等于 C_d+q^*w,那么只有 $q>q^*$ 的被告会接受,而且一旦 s^* 被拒绝,原告会用 $s^*/(C_d+q^*w)$ 的概率告下去。

那么上述猜测是否正确呢?

证明:如 $s>C_d+q^*w$,则由于 $r(s)$ 的值一定小于等于1,因此 $q(s)=[s/r(s)-C_d]/w>q^*$。根据原告的最优反应(7-C-3)式,原告应让 $r(s)=1$。

因$r(s)=1$，所以$q(s)=[s/r(s)-C_d]/w=(s-C_d)/w$，由此可知$q$值大于或等于$(s-C_d)/w$的被告会接受$s$，而小于$(s-C_d)/w$则会拒绝。

如$s\leqslant C_d+q^*w$，且$[s/r(s)-C_d]/w<q^*$，这意味着$q(s)<q^*$，根据(7－C－3)式，$r(s)=0$。但如此一来，$s=r(s)(qw+C_d)=0$，与$s>0$相矛盾，因而$q(s)$不存在(如果我们允许$s=0$，那么意味着$q(s)$可以为任意值)。

如$s\leqslant C_d+q^*w$，且$[s/r(s)-C_d]/w=q^*$，这就意味着$q(s)=q^*$，不妨设$r(s)=s/(q^*w+C_d)$，这时$q(s)$的值为$[s/r(s)-C_d]/w=q^*$，因此q值大于或等于q^*的被告会接受s^*，而q值小于q^*的被告则会拒绝。在s^*被拒绝后，原告用$r(s)=s/(q^*w+C_d)$的概率打官司。

综上所述，猜测1是诉讼博弈进行到信息集{s拒绝}的一个序列均衡。

当$s\leqslant C_d$时，均衡又如何呢？当$s\leqslant C_d$时，则会有许许多多的均衡产生。比如说，原告提出一个足够小的提议s，以至于被告肯定接受，这时信息集{s拒绝}走到的概率为0。根据序列均衡，原告在信息集{s拒绝}上的猜测可以任意设，因此不妨设原告的猜测为$f(q)$，行动为诉讼。所以，原告提出一个足够小的s，被告同意；若被告拒绝，那么原告打官司，这个策略组合就是$s\leqslant C_d$时的一个序列均衡。将$s>C_d$时和$s\leqslant C_d$时的最优审时度势组合合在一起，即得到诉讼博弈完整的序列均衡。

但$s\leqslant C_d$会产生一个问题，理论上原告要求赔偿金s可以无穷小，以致小得不合理，而要排除这种不合理，就需要一个比A2更强的假设：

A3 $E(q)w>C_p+C_d$。

在这个假设下，$s\leqslant C_d$的赔偿要求绝对不是原告的最优反应。因为$s=C_d$比任何$s<C_d$要好，而当$s=C_d$时所有可能类型的被告都会接受，因为被告无任何责任($q=0$)。而如果被告上法庭的话，他必须支付C_d，因而C_d是被告可能支付的最小值。如果$s=\bar{q}w+C_d$，那么事实上所有类型的被告都会拒绝(因为这是被告可能支付的最大值)，而拒绝之后即使原告告上法庭，他可能获得的预期收益为$E(q|s=\bar{q}w+C_d\text{被拒绝})w-C_p=E(q)w-C_p$。但由A3可知，$E(q)w-C_p>C_d$，因此原告提议$s=\bar{q}w+C_d$比提$s=C_d$要好，而提$s=C_d$又比任何$s<C_d$好，所以在A3的假设下，$s\leqslant C_d$绝对不会是原告的最优选择，因而我们可以完全不考虑$s\leqslant C_d$的情形。接下来我们要在给定猜测1成立的条件下，求解原告在开始时的最优庭外和解要求s^*是多少。

当原告提议的赔偿金额为s时，他的预期收益为$V(s)$，即

$$V(s)=s(1-F(q(s)))+F(q(s)(1-r(s))\times 0+r(s)\left[\int_0^{q(s)}qf(q)w\mathrm{d}q-F(q(s))C_p\right]$$

上式的第一项，代表被告接受和解时原告的收益；第二项代表被告拒绝和解，而原告决定不告下去的收益；第三项是被告拒绝和解，而原告告下去的预期收益。现在我们进一步假设：

A4 $\bar{q}w < 3C_p + C_d$。

那么可以证明：

(1) $V(s) < V(q^*w + C_d)$，$\forall s < q^*w + C_d$。

(2) $V'(s) < 0$，$\forall s > q^*w + C_d$。

(1) 的意思是说，如果原告要求的赔偿金比 $q^*w + C_d$ 小的话，他的预期收益一定小于 $V(q^*w + C_d)$。换言之，均衡时原告的赔偿金一定大于或等于 $q^*w + C_d$。而(2)的意思则是说，$V(s)$的值从 $q^*w + C_d$ 开始，是递减的。因此 $q^*w + C_d$ 是 $V(s)$在$[q^*w + C_d, \infty]$上的极大值。因此由(1)、(2)可知，$V(q^*w + C_d)$是 V 的极大值。我们因此得到下面的结论。

结论 在 A1、A3、A4 的假设下，整个诉讼博弈唯一的序列均衡是：原告提出 $s^* = q^*w + C_d$ 的赔偿金额。如果被告的 q 值大于 q^*，则他接受提议 s^*，否则拒绝。如果 s^* 被拒绝，原告会用 $r(s^*) = s^*/(q^*w + C_d) = 1$ 的概率打官司。

通过以上结论，我们可以分析类似于诉讼模型的均衡性质。比如，当被告或原告的诉讼成本越大时，原告所提出的赔偿金额会越大，而且原告所受到的损害值 w 与赔偿金 s 无关。又由于庭外和解概率是$(\bar{q} - q^*)/\bar{q}$，因此由 q^* 的定义可知，当损害值越高，成原告的诉讼成本越高时，和解的几率越高。

通过前面的经典举例可以看出，即使是非常简单的动态贝叶斯博弈，求解也非常不容易。而且在求解的过程中，通常还要加入一些假设，才能得到唯一的均衡结果。不过一旦求出均衡结果，通过举一反三，分析的威力就会体现出来。在阻吓定价模型中，我们可以使模型进一步复杂化，例如考虑成本、市场需求等因素对均衡行为的影响，进而判断这个模型的现实解释力。在诉讼模型中，同样可以考虑诸如诉讼成本、损害大小、不确定的程度等因素对均衡的影响，以便了解诉讼制度本身的利弊。这种分析方式构成了近年来迅速发展的法经济学。利用博弈论与经济分析的方法，通过比较各种法律制度和规则，从而研究这些制度的异同和效率性。实际上，很多法律和制度，其背后都含有效率性的考量，只是除了效率，公平公正可能更为重要，而最有效率的制度却不一定是公平的。但反过来说，如果一个制度没有效率，那么不论它是否公平，其都不会是一个好的制度，因为这时我们可以采用另外一个制度来增加某些人的福利，却不会因此而损害他人。从这个角度说，分析一个制度的效率性，依然是判别制度优劣的方法之一，而博弈论就是分析制

度形成与演变的有力工具。

㊁ 习题 7.8 请你将诉讼博弈的论文找来一读（Nalebuff，1987），看看能否读懂，最好写个读书笔记，在小组上讨论。

第四节 序列均衡的再精炼与运用

通过前面的介绍，我们已经看到序列均衡是非完全信息动态博弈的核心概念，所谓的非完全信息动态分析实际上就是求解序列均衡的过程，通过考察序列均衡的特点、性质、结构、条件，从而为各种非完全信息博弈提供解释。但是，序列均衡仍然存在令人不满意的地方。

图 7－14 所示的博弈我们在前面已经反复用到，唯一不同的地方是我们修改了里面的收益值。根据序列均衡的定义，很容易找到图 7－14 所示博弈存在一个序列均衡，即｛参与者 1 选择 R，参与者 2 在信息集｛L，M｝选择 r，$\mu_2(L)>1/2$｝。在猜测 $\mu_2(L)>1/2$ 的情况下，如果参与者 2 选择 l，那么预期收益为 $\mu_2(L)+3(1-\mu_2(L))=3-2\mu_2(L)<2$；如果选择 r，预期收益为 $2\mu_2(L)+2(1-\mu_2(L))=2$，因而参与者 2 选择 r 就是最优反应。在给定参与者 2 策略的情况下，参与者 1 选择 R 的预期收益是 3，若他背离到 L（或 M），收益为 2，因此参与者 1 选择 R 就是最优反应。

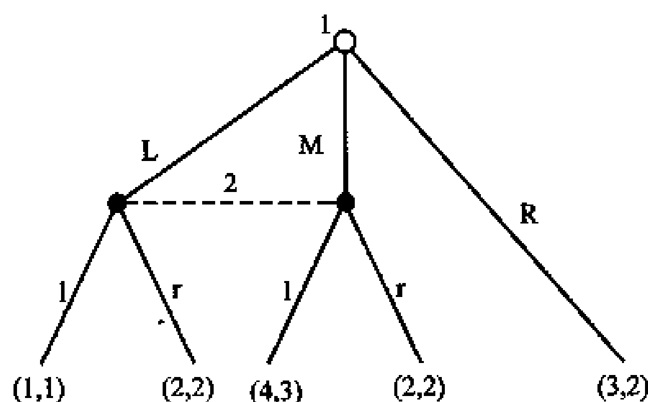


图 7－14 一个例子（3）

但是这组序列均衡却是不合理的。原因是如果参与者 1 选择 R，则他的收益为 3；如果他选择 L，则收益不是 1 就是 2。所以对参与者 1 来说，L 是一个严格劣策略（相对于 R）。早在第二章我们就已经证明过，无论在什么情况下，严格劣策略一定会被剔除，而不会成为均衡策略。因此假如博弈万一走到信息集｛L，M｝，并让参与者 2 对信息集｛L，M｝作出猜测，那么他一定会认为 $\mu_2(L)=0$。但是在上述序列均衡中，参与者 2 却猜测，一旦博弈走到信息集｛L，M｝，参与者

1 选择 L 的概率大于 1/2。因而，上述序列均衡是一个非理性的均衡。之所以出现这种矛盾，在于序列均衡对均衡路径之外的猜测(out-of-equilibrium belief)仅要求策略是给定猜测下的最优反应，而没有进一步的限制(这也是大多数情况完美贝叶斯均衡等于序列均衡的原因)。更重要的是，序列均衡完全漏掉了参与者会考虑他的对手为什么会背离原先的均衡这样的问题，因此常会产生一些不合理的均衡外猜测。在这一节里，我们主要的目的就是找出比序列均衡更合理、限制力更强的设定均衡外猜测的方法，它们最早由赵(Cho)和克瑞普斯(Kreps)在 1987 年的一篇论文中提出，这种方法又被称为序列均衡的再精炼。

一、序列均衡的再精炼

所谓序列均衡的再精炼(refinement)就是对已求解出来的均衡提出一些更强的要求和限制，从而达到强化均衡的目的，这里主要涉及三种精炼准则。为了便于读者理解，我们将利用信号博弈来加以说明。

(一) 优势准则(dominance criterion)

假设存在一个两人的信号博弈 $\Gamma=\{N,H,P,I,\rho,u\}$，其中 N 包含着自然、信号发送者(用 S 表示)、信号接收者(用 R 表示)，自然以一定的概率分布选择信号发送者的类型 t，信号发送者的类型集合为 T，即 $\forall t\in T,\rho(t)\geqslant 0$，$\sum_{t\in T}\rho(t)=1$。信号发送者知道自己的类型，但信号接收者不知道，而只能观察到信号发送者所发送的信号 m，信号发送者的信号集合为 M。不妨假设存在一个序列均衡$(\sigma_S^*,\sigma_R^*,\mu^*)$，在这个序列均衡下，R(信号接收者)收到某个信息 m 的概率为 0，称为均衡路径外信号。

定义 7.8　优势准则在信号博弈中，给定一组序列均衡$(\sigma_S^*,\sigma_R^*,\mu^*)$，对于任何均衡外信号 m，如果存在 $t\in T$ 和 $m\in M,m'\neq m$，使得

$$\min_{a\in A_R}u_S(m',a;t)>\max_{a\in A_R}u_S(m,a;t)\qquad(7-2)$$

那么一个合理的均衡应该使 $\mu^*(t|m)=0$。如果一个序列均衡符合上述限制，那它就满足优势准则。(7-2)式中的 A_R 表示信号接收者的行动空间。

这个准则非常直观：一个类型为 t 的人，如果他背离均衡而发出 m 这个均衡外信号，那么他能得到最好的收益为 $\max u_S(m,a;t)$，如果这个类型的人有另外一个信号 m' 可供选择，而他发出 m' 这个信号后，所能得到的最差收益 $\min u_S(m',a;t)$ 都还比发送信号 m 所能得到的最大收益还高，那么发信者的类型不可能为 t，因为理性的人不会做这种违反理性的事。因此如果(7-2)式成

立，那么在原先的序列均衡下，在收信者收到 m 这个均衡外信号时，他的猜测 $\mu^*(t|m)$ 一定为 0，否则用优势准则的标准来看，原来的均衡就是一个不合理的均衡，必须删除。

我们仍用图 7－14 中的序列均衡 $\{\sigma_1^*(\{\varnothing\})=\mathrm{R}, \sigma_1^*(\mathrm{L},\mathrm{M})=\mathrm{r}, \mu^*(\mathrm{L})>1/2\}$ 为例。在 $(\sigma_1^*,\sigma_2^*,\mu^*)$ 这个均衡下，L 和 M 都是均衡外信号。因为 $\min u_1(R;t)=3$ 大于 $\max\limits_{a\in\{l,r\}} u_1(L,a;t)=2$（根据序列均衡，如果参与者 1 万一走到了信息集 $\{\mathrm{L},\mathrm{M}\}$，那么参与者 2 在 $\mu_2(\mathrm{L})>1/2$ 下的最优策略是 r），所以根据优势准则 $\mu^*(\mathrm{L})=0$。因此原先的均衡设定 $\mu^*(\mathrm{L})>1/2$ 就违反了优势准则的猜测，应该从均衡集合中删除。① 对于图 7－14 中的任何一个参与者 2 选择 R 的序列均衡而言，要满足优势准则，信念 $\mu^*(\mathrm{L})$ 一定要等于 0。但在 $\mu^*(\mathrm{L})=0$ 的情况下，$\sigma_2^*(\mathrm{L},\mathrm{M})=\mathrm{l}$ 才是参与者 2 的最优反应。而给定参与者 2 的策略为 l，那么参与者 1 选择 M 就是最优反应。因而唯一一个序列均衡为：$\{\sigma_1^*(\{\varnothing\})=\mathrm{M}, \sigma_2^*(\mathrm{L},\mathrm{M})=\mathrm{l}, \mu^*(\mathrm{L})=0\}$，这是唯一可以通过优势准则的序列均衡。

优势准则是筛选合理的序列均衡的方法之一，也是最粗浅的。由于它非常直观，所以它对均衡的限制力也很弱。比如图 7－15 所示博弈中的一个序列均衡为：$\{\sigma_S^*(t)=\mathrm{u}, \forall t; \sigma_R^*(\{t_1\mathrm{u}, t_2\mathrm{u}\})=\mathrm{r}, \sigma_R^*(\{t_1\mathrm{d}, t_2\mathrm{d}\})=\mathrm{l}, \mu_R^*(t_1\mathrm{d})>1/2\}$。在这个均衡下，d 就是一个均衡路径外信号，而 u 就是一个均衡路径内信号。现在 $\min u_S(\mathrm{u}|(\sigma^*,\mu^*);t_1)=1$，而 $\max u_S(\mathrm{d}|(\sigma^*,\mu^*);t_1)=2$，由于 $1<2$，所以 (7－2) 式不成立，因而不能得到 $\mu^*(t_1|\mathrm{d})=0$ 的结论，即上述序列均衡能够通过优势准则。

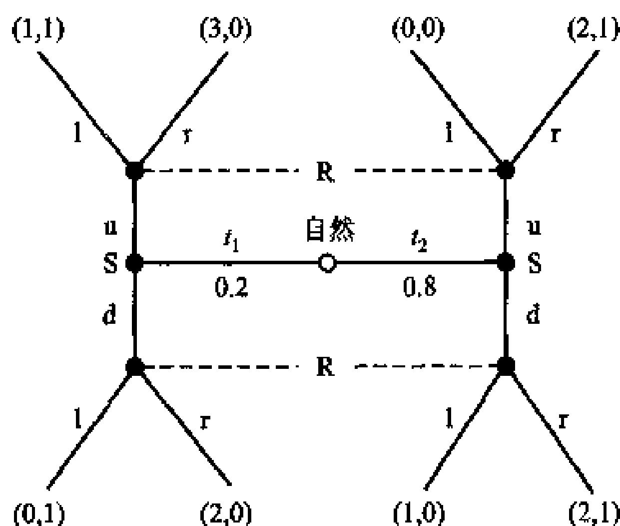


图 7－15　信号博弈与优势准则

① 在这个例子里，参与者只有一个类型，因此 $u_i(m|(\sigma^*,\mu^*);t)=u_i(m|(\sigma^*,\mu^*))$。

(易) 习题 7.9　图 7－15 所示博弈的序列均衡如上，当信号发送者类型为 t_2 时，是否符合优势准则？

优势准则的问题是，它只是用信号发送者的某一信息 m' 下的最小收益和 m 下的最大收益相比，而没有考虑到信号接收者所作的反应是否是相对于 m 和 m' 的最优反应。这表明发信者在比较发送 m 或发送 m' 之后的收益时，没有进行更深入的比较。这就引出了所谓的直观准则。

（二）直观准则（intuitive criterion）

在给出直观准则之前，我们先给出一些定义。令 $T(m)$ 为信号发送者所有可能发出 m 这个信号的类型集合，只有类型在 $T(m)$ 里的发信者，才有可能发出 m 这个信号。而 $\{\mu|\mu(T(m))=1\}$ 为信号接收者在 $T(m)$ 上的所有猜测的概率分布（密度函数）集合。设 $(\sigma_S^*,\sigma_R^*,\mu^*)$ 是一组序列均衡，而 m 是一个均衡外信号。令

$$S(m) = \{t \in T \mid u_S^*(t) > \max u_S(m,a;t),\quad a \in A_R(T,m)\} \quad (7-3)$$

其中 $u_S^*(t)$ 是类型为 t 的发信者在策略 $(\sigma_S^*,\sigma_R^*,\mu^*)$ 下得到的均衡收益，而

$$A_R(T,m) = \bigcup_{\{\mu|\mu(T(m))=1\}} A_R(\mu,T(m),m)$$
$$= \bigcup_{\{\mu|\mu(T(m))=1\}} \{a \in A \mid E_\mu u_R(m,a;t) \geq E_\mu u_R(m,a';t),\ \forall a' \in A\}。$$

其中 $\cup$ 表示并集，E_μ 表示在猜测 μ 下的期望值。

$A_R(T,m)$ 所代表的意思，是当收信者 R 收到 m 这个信号时，他会对发信者 S 的类型 t 有一个猜测 μ（这个猜测事实上是任意的[①]），在这个猜测 μ 下，R 就会有一个与 μ 相一致的最优反应 a，而 $A_R(T,m)$ 这个集合即代表当 R 收到 m 这个信号时，在所有可能对 S 的类型的猜测下，R 的所有可能的最优反应集合。

(7－3)式可以解释如下：当收信者收到 m 这个均衡外信号时，他首先会建立一个对发信者类型的猜测：$\mu(t|m)$（在序列均衡下，这个猜测可以随意设定）。与这个猜测相一致，收信者会采取相应的最优反应 a，即 $a \in A_R(T,m)$。假设发信者 S 的类型为 t，并发出信号 m，且收信者作出（在所有可能的猜测下）最优反应 $a \in A_R(T,m)$，那么 S 所有可能得到的收益中的最大收益为 $\max\limits_{a \in A_R(T,m)} u_S(m,a;t)$，如果这个最大收益也小于均衡收益 $u_S^*(t)$，那么（理性的）发信者绝对没有理由背离均衡，发出 m 这个信号。因为发信者背离均衡发出均衡外信号 m 所可能得到的最大收益仍然小于均衡收益，所以当收信者 R 收到 m 这个均衡外信号时，R 会认为发信者的类型一定不在 $S(m)$ 中。由此我们看到，$S(m)$ 所代表的是不可

① 在较一般的信号博弈里，猜测不会是任意的，因为有些类型的人根本无法发送出 m 这个信号。

能发送信号 m 的发信者的类型集合。

我们定义 $T\backslash S(m)$ 为不包含 $S(m)$ 的所有类型集合。如果发信者发出了均衡外信号 m，收信者一定相信发信者的类型不在 $S(m)$ 中，如果在 $S(m)$ 中，发信者这样做就是非理性的。但是，发信者却偏离了均衡，发出了均衡外信号 m，唯一的合理解释是：第一，发信者的类型不在 $S(m)$ 中；第二，发信者偏离均衡有利可图。这就意味着，存在某个类型 $t'\in T\backslash S(m)$，使得

$$u_S^*(t') \leqslant \min_{a\in A_R(T\backslash S(m),m)} u_S(m,a;t') \qquad (7-4)$$

成立。只有在这种情况下，S 才有偏离的动机（更高的收益）和可能（S 不可能是类型 t 的人）。上述的论述显然十分的直观，因而被称为直观准则。

定义 7.9　直观准则　在信号博弈中，给定一组序列均衡（σ_S^*，σ_R^*，μ^*），任何均衡外信号 m，如果满足(7－3)式和(7－4)式，那么一定有 $\mu^*(t|m)=0$。

直观标准具有一般性，不仅仅指信号博弈，只是信号博弈便于说明。实际上，任何一个合理的序列均衡，在任何均衡外策略 m 和类型 t 满足(7－3)和(7－4)两式时，应当有 $\mu^*(t|m)=0$。否则这个序列均衡就不符合直观准则，应该从均衡名单中删除。

接下来我们用图 7－15 来说明直观准则如何使用。

在图 7－15 所示的博弈中，我们曾给出一个能通过优势准则的序列均衡：$\{\sigma_S^*(t)=u,\forall t;\sigma_R^*(\{t_1\text{u},t_2\text{u}\})=\text{r},\sigma_R^*(\{t_1\text{d},t_2\text{d}\})=\text{l},\mu_R^*(t_1\text{d})>1/2\}$。我们现在要证明上述序列均衡不满足直观准则。在上述策略下，唯一的均衡外信号为 d，因此

$$S(d)=\{t\in T\mid u_S^*(t)>\max_{a\in A_R(T,m)} u_S(\text{d},a;t)\}=\{t_1\}$$ [①]

由此得 $T\backslash S(\text{d})=\{t_2\}$，$A_R(T\backslash S(\text{d}),\text{d})=\text{r}$，$\min\limits_{a\in A_R(T\backslash S(\text{d}),\text{d})} u_S(\text{d},a;t_2)=2\geqslant u_S^*(t_2)=2$。因此，(7－3)和(7－4)两式皆成立。根据直观准则，$\mu^*(t_1|d)=0$，而这和上述策略下的猜测 $\mu^*(t_1\text{d})>1/2$ 相矛盾，所以上述策略虽是一个序列均衡，但不符合直观准则，因而应当从均衡集合中剔除。

（三）重复直观准则（iterated intuitive criterion）

对于有些博弈，运用一次直观准则并不能把所有的非理性（均衡外策略）序列均衡都剔除掉，而是需要反复运用直观准则。当(7－3)和(7－4)两式皆成立时，收信者一定可以确定发信者的类型不在 $S(m)$ 里，所以收信者在作出他的最优反应时，应该会将 $S(m)$ 从类型集合 T 中剔除。也就是说，这时候收信者根本不需要考虑 $S(m)$ 里的类型。换言之，在(7－3)式中，收信者在作出最优反应时，考

① $u_S^*(t_1)=3$ 大于 $\max\limits_{a\in A_R(T,m)} u_S(\text{d},a;t_1)=2$，$u_S^*(t_2)=2$ 等于 $\max\limits_{a\in A_R(T,m)} u_S(\text{d},a;t_2)=2$。

虑所有的 $t \in T$ 显然不够聪明，他在选择最优反应时，事实上只需要考虑 $T\backslash S(m)$ 里的元素就够了。因此，对于每个均衡外信号 m，定义 $T^2 = T\backslash S(m)$。令

$$S^2(m) = \{t \in T^2 \mid u_S^*(t) > \max_{a \in A_R(T^2,m)} u_S(m,a;t)\}\text{①} \qquad (7-5)$$

若存在 $t' \in T^2\backslash S^2(m)$，使得

$$u_S^*(t') \leqslant \min_{a \in A_R(T^2,m)} u_S(m,a;t') \qquad (7-6)$$

那么根据直观准则，一个合理的均衡，必须令 $\mu^*(t|m) = 0, \forall t \in S^2(m)$。也就是说，这时序列均衡的猜测，不应该认为发信者 S 属于 $S^2(m)$ 里的类型。但如果满足 (7-5) 和 (7-6) 式的 $S^2(m)$ 不是空集合，那么收信者在作出最优反应时，应该把 $S^2(m)$ 从 $T\backslash S(m)$ 中剔除。只要 $S^k(m)$ 不为空集，那么这个程序就可以不断地重复进行。令 $T^k = T^{k-1}\backslash S^{k-1}(m)$，$S^k(m) = \{t \in T^k \mid u_S^*(t) \geqslant \max_{a \in A_R(T^k,m)} u_S(t,m,a)\}$。若存在 $t'' \in T^K\backslash S^K(m)$，使得 $u_S^*(t'') \leqslant \min_{a \in A_R(T^{k+1},m)} u_S(m,a;t'')$，则一个合理的序列均衡必须令 $\mu^*(t|m) = 0, \forall t \in S^k(m)$。在经过 k 次重复直观准则的步骤后，收信者对于发出信号 m 的发信者的类型的猜测范围，就会缩小到 $T^k\backslash S^k(m)$。因为 T 是有限的，因而这个过程会在有限的次数内停止，而且最后剩下的 $T^K\backslash S^K(m)$ 不会是空集。上述这个过程被称为**重复直观准则**（iterated intuitive criterion（ⅡC））。假使一个序列均衡的猜测 μ^*，能够通过第 k 次的检验，那么我们称它为一个通过 k 次重复直观准则的均衡解。而如果一均衡猜测 μ^* 能够通过所有 k 次的重复考验（请注意这个过程在有限次内要停止），那么就说它通过了重复直观准则。对重复直观准则的最形象的理解就是人与计算机下象棋的过程，在计算上显然人不如计算机来得快和准确，但人却能战胜计算机，原因就在于人们在计算如何走时，能够根据直观准则剔除明显不利的策略，从而使得自己考虑的策略集合大大缩小，而计算机却不具有像人一样的直观感觉，因而计算机相对于人而言，就是不聪明的，尽管它计算快、准确、能处理海量的数据。

二、斯彭恩的就业市场信号模型

对信号博弈的大规模深人研究始于斯彭恩（Spence，1973）提出的就业市场信号模型，该模型不仅开创了广泛运用扩展式博弈描述经济问题的先河，还较早地给出了完美贝叶斯均衡等概念。这里我们考察一个简单的就业市场信号模型，并通过它来进一步理解序列均衡的再精炼。

假设有一个雇主（R）和求职者（S）。自然决定 S 有两种可能类型，高能力

---

① $S(m)$ 和 $S^2(m)$ 之不同处，在于前者考虑所有可能的类型，而后者只考虑在 T^2 里的类型。

和低能力，$T=\{H,L\}$，高能力的先验概率 $\rho(H)=1-q$。由于雇主不能直接观察到求职者的能力，因而只能通过一个与求职者能力相关的信号来对求职者的能力作出判断。这里我们不妨假设求职者发出的信号为他受教育的程度 e，而雇主的最优回应是给出一个工资水平 w。求职者受教育程度的高低主要取决于能力的高低，能力高的人显然获得同等的教育程度付出的成本要小于能力低的人，这是斯彭思就业市场模型的一个重要假定，即

$$c_e(L,e)>c_e(H,e)。$$

其中 $c_e(t,e)$ 表示能力为 t、教育水平为 e 的求职者进一步接受教育的边际成本。根据这一假定，可知低能力的求职者的无差异曲线要比高能力的求职者的无差异曲线陡峭，如图 7-16 所示。

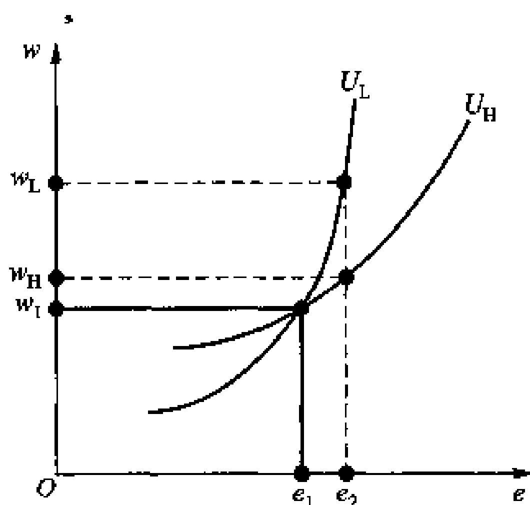


图 7-16　求职者的无差异曲线

无差异曲线表示在求职者的总效用保持不变的情况下，工资 w 和受教育程度 e 的所有可能组合。在总效用不变下，受教育程度越高，显然需要更高的工资来弥补付出的成本，所以无差异曲线斜率为正。由于低能力求职者要获得更高的教育水平需要付出更多的心血，于是就要求工资增加得更多些，才足以弥补他付出的努力。在图 7-16 中，教育水平由 e_1 增加到 e_2 时，低能力者要求工资从 w_1 增加到 w_L，而高能力者仅要求工资从 w_1 增加到 w_H。这里，我们把受教育程度解释为学生在校表现的差异，而不是受教育的时间年限。如果是受教育的年限，那么这个博弈会变成一个蜈蚣博弈，因为每一年求职者都会就是继续上学还是工作作出选择，如果选择工作，博弈结束，如果选择继续上学，那么博弈进行到下一阶段（上一节我们曾讨论过文凭的信息价值，那是另一个教育信号博弈）。

为了简单，我们假设劳动市场处于完全竞争，因而雇主给出的工资水平为一个期望值，即 $w_p(e)=\mu(H|e)\times y(H,e)+[1-\mu(H|e)]\times y(H|e)$。其中

$\mu(t|e)$表示雇主对求职者能力的猜测,而$y(t,e)$则是求职者的产出。在这样的假设下,无论求职者的策略是什么,按期望产出支付工资都是雇主的最优策略,如果雇主相信求职者为高能力者,那么$\mu(\mathrm{H}|e)=1$,从而$w=y(\mathrm{H},e)$。这样做有一个好处,就是我们可以专注研究信号发送者(求职者)的行为,而不用考虑收信者(雇主)的策略行为,同时又不失一般性。进一步的简化是假设产出$y(t,e)$为线性(实际上这个假设可有可无)。那么在信息完全的情况下,求职者可得的工资$w(e)=y(t,e)$,从而能力为t的求职者选择效用最大化的e

$$\max_e U_{\mathrm{S}}(t)=y(t,e)-c(t,e)^{①}$$

用$e^*(t)$表示最优解,如图7-17所示,工资$w^*(t)=y(t,e^*(t))$。

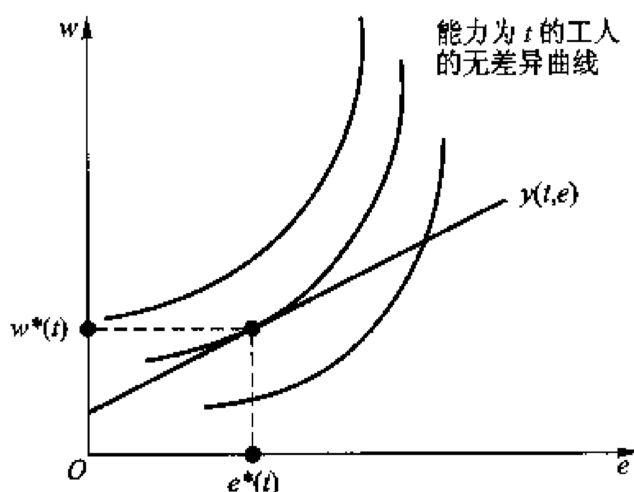


图7-17 完全信息下的工资

在开始分析之前,有一个辅助定理可以帮助我们简化推论过程。

辅助定理7.1 考虑上述就业市场模型中的某一个序列均衡,而e是这一序列均衡的均衡外信号,则不论收信者在这一均衡外信号e下的猜测是什么,这个均衡一定也可以被$\mu^*(\mathrm{L}|e)=1$这个猜测所支持。也就是说,在某一序列均衡下,在碰到任何一个均衡外信号e时,它对S的类型是L(低能力)的猜测$\mu^*(\mathrm{L}|e)$,都可以用$\mu^*(\mathrm{L}|e)=1$去代替而使均衡保持不变。

证:辅助定理7.1的证明,其实非常简单。一个序列均衡,当收信者碰到某个均衡外信号e时,他对类型L的猜测$\mu^*(\mathrm{L}|e)=1-r$,不一定等于1,这时其期望收益为$\mu(\mathrm{H}|e)y(\mathrm{H},e)+\mu(\mathrm{L}|e)y(\mathrm{L},e)=ry(\mathrm{H},e)+(1-r)y(\mathrm{L},e)$,显然当$r=0$时,上述期望收益达到最小。换言之,用$\mu^*(\mathrm{L}|e)=1$替换后,任何一个求职者偏离到$e$时,他得到的工资是最低的,用公式表示即

$$\min\{\mu(\mathrm{H}\mid e)y(\mathrm{H},e)+\mu(\mathrm{L}\mid e)y(\mathrm{L},e)\}=1\times y(\mathrm{L},e)$$

① 由于效用函数只考虑求职者,所以后面我们将省略掉下标S。

而在原先的猜测 $\mu^*(\mathrm{L}|e)$ 下，原来的审时度势已是一个均衡，这说明求职者在原先的猜测 $\mu^*(\mathrm{L}|e)$ 下，已经不愿意偏离到 e。现在如果用 $\mu^*(\mathrm{L}|e)=1$ 去取代原有的猜测，那么偏离到 e 的求职者只能得到更低的工资，因此他更不愿意偏离到 e。正因如此，用 $\mu^*(\mathrm{L}|e)=1$ 去取代原有的任何一个序列均衡的猜测，不会改变原先的审时度势组合仍旧是一个均衡这样的结果。

利用辅助定理 7.1，我们可以很容易地找出就业市场模型的分离均衡和混同均衡，更为重要的是，它可以大大简化我们的分析（只需要考虑 $\mu^*(\mathrm{L}|e)=1$ 的情况，而不用考虑 $\mu^*(\mathrm{L}|e)=1-r$ 的情况）。令 $e\hat{}(t)=\arg\max\limits_{e} y(t,e)-c(t,e)$，$e\hat{}(t)$ 是类型为 t 的求职者被认为他是类型 t 时，使他效用最大的 e。例如一个类型为 L 的求职者，当别人也知道他的类型是 L 时，令他的效用最大化的教育程度即是 $e\hat{}(\mathrm{L})$。令 $\underline{U}(\mathrm{L})=U_{\mathrm{S}}(\mathrm{L},e\hat{}(\mathrm{L}),y(\mathrm{L},e))$，表示一个低能力者讲真话所得到的效用。很明显，一个类型为 L 的求职者，在均衡时他最低的效用绝不会低于 $\underline{U}(\mathrm{L})$（因为最差不过讲真话）。令 $e\hat{}(\mathrm{H},\mathrm{L})=\arg\max\limits_{e} y(\mathrm{L},e)-c(\mathrm{H},e)$。$e\hat{}(\mathrm{H},\mathrm{L})$ 是类型 H 的求职者被人误认为他是类型 L 时（或是假装成低能力者），使他的效用达到最大的 e，这时的效用定义为 $\underline{U}(\mathrm{H})=U_{\mathrm{S}}(\mathrm{H},e\hat{}(\mathrm{H},\mathrm{L}),y(\mathrm{L},e))$。$\underline{U}(\mathrm{H})$ 代表一个类型为 H 的求职者，在均衡时所能得到的最低效用。以上叙述可以绘成图 7－18。图中越往上的无差异曲线效用越高，同时类型为 L 的求职者的无差异曲线斜率比类型为 H 者陡峭。

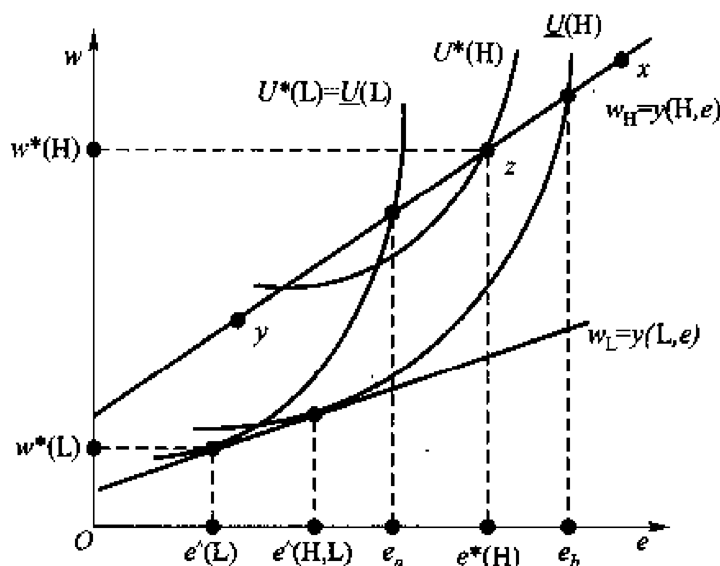


图 7－18　分离型均衡

（一）分离均衡

在分离均衡下，两种类型的求职者会发出不同的教育水平信号，因此在均衡状态下，雇主可从求职者发出的信号来判断他属于何种类型。这表明类型为 L

的求职者的工资会在 $w_L=y(L,e)$ 线上的某一点，而类型为 H 的求职者的工资会在 $w_H=y(H,e)$ 线上的某一点。一个低能力者如果要使效用最大化，那么他选择的信号一定是 $e^\wedge(L)$，即类型为 L 的求职者的无差异曲线会和 $w_L=y(L,e)$ 相切于 $e^\wedge(L)$，而他所得的效用正好等于 $\underline{U}(L)$。因此，在分离型均衡下，$e^*(L)=e^\wedge(L)$，而 $\underline{U}(L)=U^*(L)$。

接下来我们看类型为 H 的求职者。他所发出的均衡信号 $e^*(H)$，不会大于 e_b。这是因为如果 $e^*(H)>e_b$，那么类型为 H 的求职者得到的工资会在 $w_H=y(H,e)$ 线上 e_b 点之右的某一点上（例如图 7－18 中的 x 点）。但如果这样，他宁可发出 $e^\wedge(H,L)$ 这个信号。这是因为如果他发出 $e^\wedge(H,L)$，那么根据辅助定理 7.1，他会被认为是类型为 L 的求职者，得到 $w=y(L,e)$ 的工资，因此得到 $\underline{U}(H)$ 的效用。而由图 7－18 中可以看出，通过 x 点的无异曲线的效用要小于 $\underline{U}(H)$。因此偏离到 $e^\wedge(H,L)$ 所得到的效用会比发出 $e^*(H)>e_b$ 的效用要高。

$e^*(H)$ 也不会小于 e_a。因为如果 $e^*(H)<e_a$，那么他会得到像图 7－18 中 y 点的工资。另一方面，在 y 点，类型为 L 的求职者所得到的效用大于 $U^*(L)$，因此他一定会偏离 $e^*(L)$，而发出 $e^*(H)$ 这个信号，这样他会被认为是类型为 H 的求职者，从而得到更高的效用。这和 $e^*(L)$ 是均衡信号的结论相矛盾。因此 $e_a\leqslant e^*(H)\leqslant e_b$。

命题 7.1 只要任何一个 $e^*(H)$ 满足 $e_a\leqslant e^*(H)\leqslant e_b$，那么下列审时度势组合是分离均衡：

(1) 求职者的策略：$e^*(L)=e^\wedge(L)$，$e^*(H)=e^\wedge(H)$。

(2) 雇主的信念：$\mu^*(L|e^*(L))=1$；$\mu^*(H|e^*(H))=1$；$\mu^*(L|e)=1$，$\forall e\neq e^*(L)$ 和 $e^*(H)$。

(3) 雇主的策略：如果高能力，那么 $w^*=y(H,e^*(H))$；如果低能力，那么 $w^*=y(L,e^*(L))$。

证明：由于分离型均衡的特性，再加上辅助定理 7.1 的结论，命题 7.1 中的条件(2)对 μ^* 的设定明显成立。因此我们只要证明在 μ^* 的设定下，$e^*(L)$ 和 $e^*(H)$ 各自是求职者的最优回应即可。如果类型为 L 的求职者没有发出 $e(L)=e^\wedge(L)$，而发出 $e^*(H)$，那么他会被认为是类型为 H 的求职者，而得到的工资会在 $w_H=y(H,e)$ 这条线上，且对应于 $e^*(H)$ 的 z 点上。但由于 $e^*\geqslant e_a$，因此由图 7－18 可以很容易看出，类型为 L 的求职者在这点所得到的效用反而变低了。因此类型 L 的求职者不会发出 $e^*(H)$ 的信号。如果他发出任何不等于 $e^*(H)$ 的讯号，那么由 μ^* 的设定，他会被认为是类型为 L 的求职者。但由 $e^\wedge(L)$ 的定义可知，$e^\wedge(L)$ 已经是类型 L 的最优信号，因此类型为 L 的求职者发出 $e^\wedge(L)$

是他的最优策略。接着我们要证明 e^*(H)是类型为 H 的求职者的最优策略。如果他偏离到任何一个不等于 e^*(H)的信号,那么由 μ^* 的设定,他会被认为是类型 L 的求职者,因而所得到的工资会在 $w_L = y(L,e)$ 这条线上。由图 7-18 可以看出,这样的效用绝对不可能比 U^*(H)大。这就证明了命题 7.1。

（二）混同均衡

在混同型均衡下,类型为 L 的求职者和类型为 H 的求职者会发出一样的信号,因而得到一样的工资,令 $e^*(L) = e^*(H) = e^*$。由于两个求职者的信号相同,所以雇主对求职者的猜测仍是原先的先验猜测,也就是 $\mu(L) = q$。由于求职者的工资是一个均值,因此在均衡状态下,两个求职者的工资在 $w_p = \mu(L)y(L) + \mu(H)y(H)$ 这条线上(见图 7-19)。

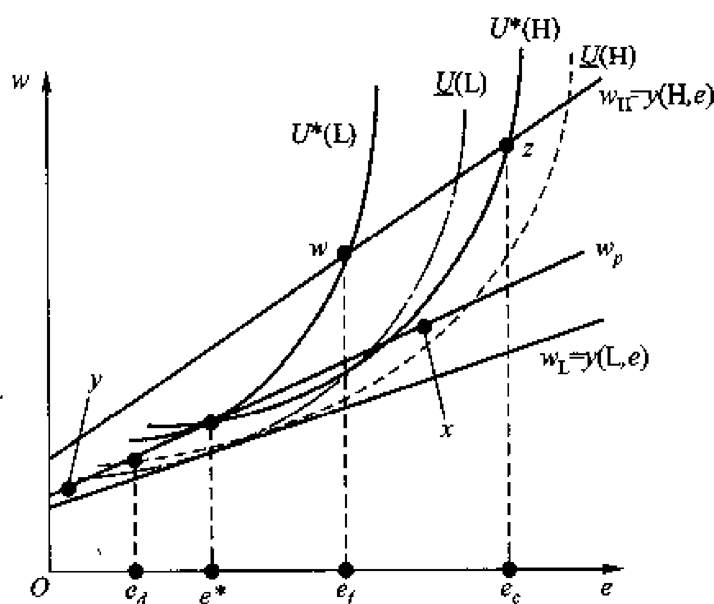


图 7-19 混同型均衡

我们要证明,在均衡的时候,两个求职者所发出的信号,一定位于 e_d 和 e_f 之间。这个原因很简单。如果 $e^* > e_f$,那么类型为 L 的发信者所得到的效用,会比 U^*(L)还要低。比如,无差异曲线经过图 7-19 中的 x 点的类型为 L 的求职者,所得到的效用会在 U^*(L)这条无异曲线之下,因此类型为 L 的求职者愿意偏离 e^* 而发出 $e^{\wedge}$(L)的信号,以得到 U^*(L)的效用。如果 $e^* < e_d$,那么类型为 H 的发信者所得到的效用,会比 $\underline{U}$(H)还要低。例如,经过图 7-19 中 y 点的类型为 H 的求职者,所得到的效用会在 $\underline{U}$(H)这条无异曲线之下,因此类型为 H 的求职者宁可偏离 e^*(H)而送出 $e^{\wedge}$(H,L)的信号,以得到 $\underline{U}$(H)的效用。不但如此,我们还可以证明在 e_d 和 e_f 之间的任何信号 e^*,都可以是混同均衡。

命题 7.2 只要 $e_f \geq e^* \geq e_d$,那么下列审时度势组合是混同均衡:

(1) 求职者的策略：$e^*(L)=e^*(H)=e^*$。

(2) 雇主的信念：$\mu^*(L|e^*)=q,\mu^*(H|e^*)=1-q;\mu(L|e)=1,\forall e\neq e^*$。

(3) 雇主的策略：给出的工资 $w_p=\mu(L)y(L)+\mu(H)y(H)$。

证明：(2)的证明很容易，请读者自己证明。证明(1)等价于证明 $e^*(L)$ 和 $e^*(H)$ 各是求职者的最优反应。定义 $\max U^*(t)=w_p-c(t,e)$。如果类型为 L 的求职者发出任何一个 $e\neq e^*$ 的信号，那么由(2)的设定可知他会被认为是类型为 L 的求职者。但由于 $e_f\geqslant e^*$，因此很容易可以由图 7－19 得出 $U^*(L)>\underline{U}(L)$，由此可知 $e^*(L)=e^*$ 是类型为 L 的求职者的最优反应。如果类型为 H 的求职者发出任何一个的 $e\neq e^*$ 信号，那么由(2)可知，他会被认为是类型为 L 的求职者。因此他最好也只能得到 $\underline{U}(H)$ 的效用。但由于 $e^*\geqslant e_d$，因此由图7－19很容易看出 $U^*(H)>\underline{U}(H)$，由此可知 $e^*(H)=e^*$ 是类型为 H 的求职者的最优反应。

命题 7.1 和命题 7.2 所找到的均衡其实并不是所有的均衡，因为还有很多均衡既不是混同型也不是分离型，但那些均衡一方面求解不易，另一方面也不太有直觉上的意义，因此我们只对上述两类的均衡有兴趣。接下来我们要看看，是不是所有的上述的均衡，都可以通过直观准则。

先看混同型均衡。类型为 L 的求职者，如果发出一个在 e_f 和 e_c 之间的信号的话，得到的工资最多也只会在 $w_H=y(H,e)$ 这条线上的 w 和 z 之间的点上(见图 7－19)。但类型为 L 的求职者在这里任何一点的效用，都低于 $U^*(L)$。因此对于任何 $e'\in(e_f,e_c)$，$L\in S(e')=\{t\in T|U^*(t)\geqslant\max\limits_{w\in A_R(T,e')}U_S(t,e',w)\}$，其中 $T=\{H,L\}$。这表明，类型为 L 的求职者根本不会送出 e'这个信号。因此，如果雇主收到 e'，他会认为求职者的类型是 H，而给他$\overline{wz}$这条线段上对应于 e'的工资。一旦如此，类型为 H 的求职者的效用会大于均衡效用 $U^*(H)$，所以一定有动机偏离到 e'，以证明自己是类型为 H 的求职者，从而得到更高的效用，即 $\min\limits_{w\in A(T\backslash L,e')}U^*(H,e',w)\geqslant U^*(H)$，因此(7－3)和(7－4)式成立。根据直观准则，我们有 $\mu^*(L|e')=0$。因此原先的所有混同型均衡中的猜测，$\mu^*(L|e')$必须等于0。一旦如此，类型为 H 的求职者一定会偏离原先的均衡信号 e^*，而发出 $e'\in(e_f,e_c)$这个信号。所以类型为 H 的求职者一定可以找到一个 $e'\in(e_f,e_c)$使得原先的均衡不能满足直观准则，由于 e^* 可以为 e_d 和 e_f 之间的任意数，所以没有一个混同均衡能够通过直观准则。

接下来我们再看看分离均衡是否可以通过直观准则。令 $e\in(e_a,e_b)$。一个类型为 L 的求职者，如果发出 e 这个信号，那么即使他得到类型为 H 的求职者的工资，他的效用也会低于 $\underline{U}(L)$(见图 7－18)。因此 $\underline{U}(L)>\max\limits_{w\in A_R(T,e)}U^*(L,e,w)$。这

表明 $S(e)=\{L\}$。但由于 $T\backslash S(e)=\{H\}$，且由图 7－18 可得 $U^*(H)<\min_{w\in A_R(T\backslash S(e),w)}U^*(H,e,w)$。因此由直观准则可知 $\mu^*(L|e)=0$。这意味着，类型为 L 的求职者如果发出 e 这个均衡外信号，一定会使自己的效用水平变低，因此雇主会认为类型为 L 的求职者不会发出 e 这个信号，即 $S(e)=\{L\}$。但如果雇主真这么认为，那么在收到 e 后就会给求职者类型为 H 的高工资，那么类型为 H 的求职者就愿意偏离到 $e(<e^*(H))$ 这个均衡外信号，从而使新的效用高于原来的均衡效用 $U^*(H)$。所以直观准则要求当雇主收到 e 这个信号时，应该把他对求职者是属于 L 的猜测调为 0，亦即 $\mu^*(L|e)=0$。换言之，在原先所有的分离型均衡里，我们必须让 $\mu^*(L|e)=0$。但如果这样，类型为 H 的求职者一定会偏离 $e^*(H)$ 而发出 $e\in(e_a,e^*(H))$ 的信号，因此原先的均衡就不再是均衡。这个情形的产生，完全是因为 e_a 和 $e^*(H)$ 之间的信号都具有上述特性。所以要使类型为 H 的求职者不会偏离 $e^*(H)$，唯一的可能是令 $e^*(H)=e_a$。所以 $e^*(H)=e_a$ 是分高均衡下，类型为 H 的求职者唯一可能发出的信号。

综上所述可知，在就业市场信号模型中，虽然存在着许许多多的分离型、混同型、半混同型、杂合型的均衡，但除了唯一一个分离均衡能通过直观准则的检验外，其余的序列均衡没有一个能满足直观准则。直观准则的使用，剔除了几乎所有的序列均衡，而只留下对均衡的唯一一个预测。

㊥ 习题 7.10　利用图 7－15 所示博弈（如果必要可以修改其中的收益值），讨论直观准则的运用。

㊇ 习题 7.11　阅读斯彭斯 1973 年发表的就业市场论文，并作读书笔记，在课堂上讨论。

第五节　均衡概念小结

人类社会的一切矛盾实质上都根源于人们之间的经济利益冲突，人类社会的各种制度的建立，大到上层建筑，小到单位的规章，实际上都是为了解决人类社会的各种冲突而规定的一种游戏规则。从经济学本身来看，脱离社会关系的经济学必然是残缺和无意义的，然而要克服经济学的这种残缺，除了勇气和毅力，还需要达成目标的手段，而博弈论就是这样一种手段，它使经济学变得丰满了起来，使经济学鲜活了起来，也使人类理解自身的社会性有了深人的可能。亚里士多德早在几千年前就曾说过："人是社会性动物"，而博弈论就是我们打开人的社会性大门的钥匙。也许这是对学习博弈论有什么用最好的回答。

本节中，我们对所学习的博弈均衡概念作一个简单的总结。从博弈论中的

均衡概念可以看出，序列均衡是最一般的均衡概念，它包含着所有其他均衡概念，子博弈完美均衡、贝叶斯纳什均衡和纳什均衡都是序列均衡分别在动态完全信息博弈、静态非完全信息博弈和完全信息静态博弈下的特例。一个序列均衡一定是子博弈完美均衡、贝叶斯纳什均衡和纳什均衡，但反之不成立。从纳什均衡向上走，每一个均衡概念限制的条件会越来越强，子博弈完美均衡是一个剔除不可置信威胁的纳什均衡，而贝叶斯纳什均衡则是在理性猜测下的纳什均衡，到了序列均衡不仅要求剔除不可置信威胁，而且要求更强的猜测设定，所有策略都是在给定猜测下的最优策略。从理论上讲，运用序列均衡概念几乎可以分析所有的社会经济现象，但令人遗憾的是，至今求解序列均衡仍没有一般可用的方法，哪怕一个简单的模型求解序列均衡也十分烦琐，通常需要我们先去猜出可能的均衡，然后再加以证明。尽管如此，博弈论仍然为经济学的发展带来了一场真正意义上的革命。通过博弈论提供的理论，制度的演化成了经济学研究的对象。通过人工智能模拟各种可能的演化机制，为法经济学、行为经济学与实验经济学等的研究提供了基本的条件，而这种研究的新方向构成了今天经济学最为激动人心的领域。

总而言之，掌握博弈论已成为从事经济学研究必备的基本素质。

㊣ 习题 7.12　运用所学的博弈论知识，写一篇小论文，将你在生活中观察到的现象概括为一个模型，并作出分析与解释。

关键词

自然　信息集和信息分割　博彩　贝叶斯扩展式博弈 $\Gamma=\{N,H,P,I,\rho,u\}$　泛子博弈　信念系统　行为策略　完美贝叶斯均衡　序列均衡　序列理性　信念与策略的一致性　信号博弈　分离均衡　混同均衡　半混同均衡　杂合均衡　连锁店悖论　序列均衡的再精炼　优势准则　直观准则　重复直观准则

内容要点

1. 动态贝叶斯博弈的扩展式是一般的扩展式，策略式博弈、完全信息扩展式博弈、静态贝叶斯博弈都是一般扩展式的特例，一般扩展式与完全信息扩展式博弈相比，最大的差异就在于对信息集的处理，不同的信息集结构（即信息分割）表达着不同的信息不完全。对于信息集，通常要求满足三个特性：信息集中的每一个子历史之后轮到同一个参与者行动、行动空间完全一样、信息分割。

2. 在动态贝叶斯博弈中简单地套用子博弈完美均衡会产生非常不合理的

结果，根本原因在于子博弈完美均衡没有办法处理信息集走到概率为0的情况，由此引出了完美贝叶斯均衡的概念，但完美贝叶斯均衡在发生概率为0的信息集随意设定猜测仍然是不合理的，引入序列策略和序列信念以克服完美贝叶斯均衡随意设定猜测的缺点，就形成了序列均衡。不过幸运的是，在大多数情况下，完美贝叶斯均衡与序列均衡是等价的。

3. 在动态贝叶斯博弈中，参与者的策略离不开信念，信念也离不开策略，策略与信念必须一致，策略+信念就构成了参与者真正的“万全之策”，又被称为审时度势。这说明，在贝叶斯博弈中，信念的作用被突出了出来，在完全信息博弈中，信念的作用是隐含的。实际上，对信息完全博弈与信息不完全博弈分析的差异，关键就在于对信念是如何处理的。

4. 在有些情况下，序列均衡仍然不够严格，可能包含着多个均衡，而其中有些均衡仍然不合理，这就引出了对序列均衡的再精炼。序列均衡的再精炼涉及三个原则：优势准则、直观准则、重复直观准则。

进一步阅读

Kreps 和 Wilson(1982)的论文最早给出了序列均衡的概念，而对于序列均衡的再精炼则可以参考 Cho 和 Kreps(1987)的论文，Spence(1973)的论文开创了信号博弈的研究，也为信息经济学的研究奠定了一个重要的模型。在 Osborne (2004)的论文中，给出了一些经典的动态贝叶斯博弈例子，非常通俗易懂。Mas-Colell、Whinston 和 Green(1995)则对完美贝叶斯均衡和序列均衡给出了一个非常直观且严格的说明。

第八章

若干专题及扩展

本书的最后一章将对纳什均衡的一些扩展进行一个介绍性的讨论，这些扩展反映着博弈论最前沿的研究与最新的运用。同时，这些扩展也为更好地理解博弈论提供了新的角度和基础。这一章主要涉及三个专题，即同盟博弈及其解概念——核、演化博弈及其解概念——演化稳定策略和学习博弈，之所以选择这三个专题，是因为现有的博弈论教材对这三个专题介绍较少，而实际上它们是非常重要的。

第一节　同盟博弈与核

同盟博弈(coalitional game)所关注的是博弈参与者是如何组成各种利益集团的，博弈参与者唯一能做的就是站队，一旦选择加入某个利益集团，参与者的收益也就被决定。我们把一个利益集团称为一个**同盟**(coalition)，如果一个同盟包含所有的博弈参与者，就称为**大同盟**(grand coalition)。实际上，如果我们把所有博弈参与者看作一个全集，用 N 表示，那么一个同盟就是 N 的子集，并且是对 N 的一个**分割**(partition)，即如果我们定义一个博弈中的各同盟为 S_i，$i=1,\cdots,k$，那么 $S_i \cap S_j = \varnothing$，$i \neq j$ 和 $\cup_{i=1}^{k} S_i = N$。为了方便，我们用 N 来表示大同盟，用 S 来表示其他的任意同盟。显然，大同盟是一个极端，而每一个参与者都是一个同盟是另一个极端，而大多数同盟分割都位于两个极端之间。对于同盟博弈而言，主要感兴趣的是位于两个极端之间的同盟分割情况。由于博弈参与者唯一能做的就是站队，因而参与者的偏好实际上就是对各利益集团(同盟)的行动进行排序。

定义 8.1　**同盟博弈**可表示为 $G=\{N,A_S,u\}$，其中：

(1) 参与者集合：$N=\{1,\cdots,n\}$。

(2) 对于每一个同盟 S 而言的行动集合：A_S。

(3) 偏好：每一个参与者的偏好就是对(其参与其中的)各同盟行动的排序。

【例8.1】 景泰蓝的分工协作

要生产出一个景泰蓝需要 n 个劳动者分工协作，生产出来的景泰蓝所获收益（不妨假设为1）在各工人之间进行分配。只要有一个工人不合作，景泰蓝就无法生产出来。每一个工人都希望自己获得的收益份额越多越好。该情景的同盟博弈为 $G=\{N,A_S,u\}$：

(1) 参与者集合：$N=\{1,\cdots,n\}$。

(2) 行动集合：只要有一个工人不合作，那么每一个工人所获收益为0；对于大同盟 $N=\{1,\cdots,n\}$ 的行动集合为所有向量 $\{a_1,\cdots,a_n\}$ 的集合，其中 $a_1+\cdots+a_n=1, a_i>0, i\in N$。

(3) 偏好：每一个参与者希望获得的收益越多越好。

【例8.2】 地主与农工

一个地主的土地如果由 k 个农工耕种的话，能生产出 $f(k+1)$ 的产量，其中 f 为单调增函数，$f(0)=0$。农工数为 $m\geq k$。地主与每一位农工都希望获得的产量份额越多越好。该情景的同盟博弈为 $G=\{N,A_S,u\}$：

(1) 参与者集合：$N=\{$地主，m 个农工$\}$。

(2) 行动集合：纯由农工组成的同盟有唯一的一个行动，其中每一个成员的收益为0；由地主和 k 个农工组成的同盟的行动集合则可表示为产出 $f(k+1)$ 的所有可能分配份额组合集合。

(3) 偏好：获得的收益份额越高越好。

通常我们将同盟 S 中成员的收益分配份额组合称为 S—配置，如果是 N—配置就简称为配置。

【例8.3】 多数表决

三个人对一个单位为1的产出进行表决。任何多数——两人或三人——将决定该产出如何进行分配。每一个人都希望自己获得的份额最大。该情景的同盟博弈为 $G=\{N,A_S,u\}$：

(1) 参与者集合：$N=\{1,2,3\}$。

(2) 行动集合：只包含一个人的同盟有唯一一个行动，收益为0；由两人或是三人组成的同盟，其行动集合则为所有可能的 S—配置集合。

(3) 偏好：获得的份额越高越好。

上面这些例子中，每一个同盟 S 的行动集合实际上就是 S 所获收益的 S—配置集合，S 中成员的偏好由其所获的收益份额代表。因此，我们可以用一个数来表示同盟 S 的行动集合，它实际上就是同盟 S 所获得的总收益。具有这种特征的同盟博弈就被称为转移支付（transferable payoff）同盟博弈。

定义8.2 转移支付同盟博弈 如果存在一个收益函数集合，其中每一个

收益函数都代表一个参与者的偏好，并且对每一个同盟 S，S 的每一个行动就是对同盟 S 的所获总收益在同盟成员间的一个收益分配组合。具有这个特征的同盟博弈就是所谓的转移支付同盟博弈。

我们把转移支付同盟博弈中的同盟 S 的总收益称为 S 的价值，定义为 $v(S)$。这样，一个转移支付同盟博弈就可以用参与者集合及其价值函数表示。

例如，景泰蓝分工协作博弈中：$N=\{1,\cdots,n\}$，$v(1)=v(2)=\cdots=v(n)=0$，$v(\{1,\cdots,2\})=1$。地主和农工博弈中：$N=\{0,1,\cdots,m\}$（0 代表地主），$v(0\notin S)=0$，$v(0\in S)=f(k+1)$。在多数表决博弈中：$N=\{1,2,3\}$，$v(1)=v(2)=v(3)=0$，$v(\{1,2\})=v(\{1,3\})=v(\{2,3\})=v(\{1,2,3\})=1$。

同盟博弈的目的在于表明参与者结成同盟要比单打独斗好，如果大同盟能够带给每一个参与者至少像任意小同盟（即 N 的任意子集）一样好的结果，那么这个博弈就是具有凝聚力的（cohesive）。

定义 8.3　具有凝聚力的同盟博弈　一个同盟博弈是具有凝聚力的，如果对于大同盟 N 的任意分割 $\{S_1,\cdots,S_k\}$ 和每一个行动组合 $(a_{S_1},\cdots,a_{S_k})$（其中 a_{S_j} 为同盟 S_j 的行动），对于每一个参与者 i 而言，大同盟存在一个行动至少与任一同盟 S_j 的行动 a_{S_j} 一样好，其中 $i\in S_j$。

从上述定义很容易得出，当且仅当对于 N 的任一分割 $\{S_1,\cdots,S_k\}$，$v(N)\geqslant v(S_1)+\cdots+v(S_k)$，转移支付同盟博弈是具有凝聚力的。

在同盟博弈中，最核心的概念就是所谓的核（core）。如果不存在一个同盟使得其成员偏好它的行动胜于大同盟的行动，那么大同盟的行动就是"稳定的"，大同盟所有稳定的行动集合就是所谓的核。其正式的定义如下：

定义 8.4　同盟博弈的核　一个同盟博弈的核是大同盟 N 的行动 a_N 的集合，以至于对于任一同盟 S 中的成员而言，偏好 a_N 胜于 a_S。

如果一个同盟 S 存在一个行动 a_S，使得其所有成员都偏好 a_S 胜于大同盟 N 中的某个行动 a_N，那么我们就认为 S **能够改善** a_N。因而，一个核又可以定义为不存在任一同盟能够改善行动 a_N。由于核是一个行动集合，因而核始终是存在的（空集也是一个集合）。由于我们假设同盟中的各成员只关心自己所在同盟的行动，因而一个同盟的成员考虑背离时，根本不用考虑余下成员的行为。

在一个转移支付同盟博弈中，当且仅当一个同盟 S 的价值 $v(S)$ 超过同盟 S 中的成员在 a_N 中的相应收益之和，一个同盟 S 能够改善大同盟的行动 a_N。正式的表述如下：

命题 8.1　a_N 属于转移支付同盟博弈的核，当且仅当对于每一个同盟 S，S 的成员在大同盟 N 的行动 a_N 中的总收益 $x_S(a_N)$ 不小于 $v(S)$，即

$$x_S(a_N)\geqslant v(S)$$

对每一个同盟 S 成立。

利用命题 8.1 可以对同盟博弈的 a_N 是否为核进行验证，从而为求解同盟博弈的核提供了一种方法。

【例 8.4】 地主与两李

地主的一块地由大李或小李耕种，其产量为 $f(2)$；如果由两人同时耕种，其产量为 $f(3)$。如果没有土地显然不可能有产出；如果没有农工，只有土地，那么庄稼会自然生长，产量为 $f(1)$。令 (x_1,x_2,x_3) 为大同盟的一个行动，即产出 $f(3)$ 在三人之间的一个配置。因此，(x_1,x_2,x_3) 属于该同盟博弈的核，当且仅当

$$x_1 \geqslant f(1)\text{（即地主组成的同盟）}$$

$$x_2 \geqslant 0\text{（即大李组成的同盟）}$$

$$x_3 \geqslant 0\text{（即小李组或的同盟）}$$

$$x_1 + x_2 \geqslant f(2)\text{（即由地主和大李组成的同盟）}$$

$$x_1 + x_3 \geqslant f(2)\text{（即由地主和小李组成的同盟）}$$

$$x_1 + x_2 + x_3 = f(3)\text{（即必须是一个配置）}$$

利用最后一个条件，可知 $x_1 = f(3) - x_2 - x_3$，上述条件又可重新写作：

$$0 \leqslant x_2 \leqslant f(3) - f(2)$$

$$0 \leqslant x_3 \leqslant f(3) - f(2)$$

$$x_2 + x_3 \leqslant f(3) - f(1)$$

$$x_1 + x_2 + x_3 = f(3)$$

由此可知，如果行动 (x_1,x_2,x_3) 在核中，那么每一个农工获得的最大收益不会超过 $f(3) - f(2)$（即第三个人参加到两人组成的同盟中所带来的额外收益），而两个农工所获得的最大收益不会超过 $f(3) - f(1)$（即两个农工参加到地主组成的同盟中所带来的额外收益）。满足这两个条件的行动 a_N 就是地主与两李的核，同时由此形成的大同盟才会具有凝聚力，也才是稳定的。实际上，利用上面分析得出的结论，很容易就能得出例 8.2 地主与农工中的核。

㊃ 习题 8.1　请你给出例 8.2 地主与农工同盟博弈的核。

㊃ 习题 8.2　证明多数表决的核为空集。

通过上面的介绍，可以看出同盟博弈研究的核心就在于探讨利益集团的稳定性条件，因而核所包含的思想与纳什均衡完全一样，可以看作是对纳什均衡的一种扩展。通过运用同盟博弈，我们可以分析各种各样的利益集团，大到军事、政治同盟，小到几人组成的小帮小派。无论利益集团的形式如何多变，如何在成员间进行利益分配始终是利益集团的核心问题，利益集团的崩溃通常都是由于成员内部分配不均所导致。

另外，同盟博弈还给了我们一些有益的启示。如果利益集团争夺的总收益不变，那么通常核为空集，这时各参与者之间的利益是冲突的，此消彼长，形象的比喻就是俄罗斯不可能加入美日军事同盟中。当总收益能够增加，那么核通常不为空，因而要实现一个社会的和谐，就需要能够增加社会总收益这个大蛋糕，所以说发展是硬道理。

第二节 演化博弈与演化稳定策略

根据达尔文的进化理论，生物的一种行为方式能够被保留下来通常是由于在繁衍后代上不存在其他更为有效的方式。例如，草原上的野兔一出生就知道四处张望以防不测，原因就在于不这样做的野兔都被消灭光了。在一个生物体的行为彼此相互依存的环境中，一种行为方式的成功复制（在下一代身上）依赖于其他生物体的行为。例如，在一个所有生物都采用奔跑以求生存的环境中，一个生物（变种）采用原地躲藏可能将难以生存；相反，在一个以躲藏为主的生物群落中，如果采用奔跑可能就是自我暴露，首先被吃掉。显然，博弈论为研究这种环境下的演化提供了可能的工具。另一方面，在博弈论研究中，我们通常假设参与者拥有足够的推理能力且充分理性，而实际上，生物的进化可能是基于简单的原则，通过竞争的方式来淘汰掉任何非最优的行为。这就意味着，在很多情况下，演化的方法能为纳什均衡及其精炼提供新的基础。

演化从本质上来讲是一个动态过程，为了简单，我们只考虑静态的情况。此外，我们首先从最简单的模型开始——就是所谓的**单态纯策略模型**。所谓单态，就是种群中的所有生物（除了变种）都遵从完全一样的行为方式。同时，进一步假设存在一个生物种群，其成员随机两两配对（不妨假设随机抽取两个生物），每个生物的行为集合完全一样，一个生物以一个较高的概率继承父母的行为方式，而以一个较低的概率选择一种变异行为方式，收益就是生物的适应性（即健康后代的成活数）。因而，上述演化模型就可以用一个行动有限的两参与者对称策略博弈来表述。

那么，对于任意一个行动有限的两参与者对称策略博弈，一个自然的问题就是在什么样的环境下，行动 a^* 是演化稳定的？例如，在一个生物种群中，出现了若干选择行动不为 a^* 的变种，如果这个生物种群是稳定的，那么变种所获得的收益就应当小于正常生物的收益（记住收益是繁衍的后代数），以至于变种最后会死光（如果收益一样，那么正常生物与变种在种群中的比例不变，这也可以看作是一种稳定状态，但为了简单我们不作考虑）。不妨定义 ε 为种群中变种生物所占的比重，$1-\varepsilon$ 就为种群中正常生物的比重。设变种采取的行动为 b。在

随机的配对中，变种遇到采用行动 a^* 的正常生物的概率等于 $1-\varepsilon$，遇到采取行动为 b 的变种的概率为 ε。因而，一个变种的期望收益就为

$$(1-\varepsilon)u(b,a^*)+\varepsilon u(b,b)$$

相似的，一个正常生物的期望收益就为

$$(1-\varepsilon)u(a^*,a^*)+\varepsilon u(a^*,b)$$

其中 $u(\cdot)$ 为收益函数，我们约定俗成规定当事者的行动始终排在第一，而第二个行动为对手的行动。

如果该生物种群要保持稳定状态，即行动 a^* 是演化稳定的，那么变种在种群中的数量就应当逐渐减少以致消失，这意味着正常生物的收益要大于变种的收益，即

$$(1-\varepsilon)u(b,a^*)+\varepsilon u(b,b)<(1-\varepsilon)u(a^*,a^*)+\varepsilon u(a^*,b) \tag{8-1}$$

对所有的 $b\neq a^*$ 成立。

由于式(8-1)中 ε 的值不确定，因而上述不等式要成立实际上等价于是否存在一个 $\bar{\varepsilon}>0$，使得当 $0<\varepsilon<\bar{\varepsilon}$ 时不等式成立。这就有了演化稳定性的定义：

定义 8.5 行动 a^* 是**演化稳定的**，假如存在一个 $\bar{\varepsilon}>0$，使得当 $0<\varepsilon<\bar{\varepsilon}$ 时，不等式(8-1)成立。

根据定义 8.5 来证明一个行动是否是演化稳定的将有一点烦琐，如果能够避免寻找 $\bar{\varepsilon}$，就能起到简化的作用。

首先，我们有如下结论：

如果存在一个 $\bar{\varepsilon}$，使得对所有的 $0<\varepsilon<\bar{\varepsilon}$，$a^*$ 满足(8-1)式，那么 (a^*,a^*) 就是一个纳什均衡。

㊐ *习题 8.3 请你证明上述结论。(用反证法。)*

其次，假如 (a^*,a^*) 是严格纳什均衡，那么就存在一个 $\bar{\varepsilon}>0$，以至于对所有的 $\varepsilon\in(0,\bar{\varepsilon})$，$a^*$ 满足(8-1)式。

㊐ *习题 8.4 请你证明上述结论。*

将上述两个结论合并，就得到有关演化稳定性的新定义。

定义 8.6 策略博弈中的演化稳定行动(纯策略) 在一个两参与者对称策略博弈中，一个参与者的行动 a^* 是演化稳定的，如果它满足下列条件：

(1) (a^*,a^*) 是一个纳什均衡；

(2) $u(b,b)<u(a^*,b)$ 对于每一个 $b\neq a^*$ 成立，其中 b 是 a^* 的一个最优反应。

其中 $u(\cdot)$ 中的第一个行动为当事者的行动，第二个行动为对手的行动，后面的也如此，不再重复。

需要注意的是，如果(a^*,a^*)是严格纳什均衡，那么第二个条件就是多余的。定义8.6既指出了演化稳定与纳什均衡的相同点，也指出了它们的不同。根据定义8.6易知，演化稳定行动组合必是纳什均衡，但纳什均衡不一定是演化稳定的。

考虑如图8－1所示博弈。假设正常生物选择行动X，而变种则选择行动Y，变种在整个种群中的比重为ε，因而一个生物遇到变种的概率为ε。

	X	Y
X	3,3	0,0
Y	0,0	1,1

图8－1　生物配对博弈矩阵(1)

可知，一个正常生物的期望收益为

$$3(1-\varepsilon)+0\cdot\varepsilon=3(1-\varepsilon)$$

一个变种的期望收益为

$$0\cdot(1-\varepsilon)+1\cdot\varepsilon=\varepsilon$$

根据定义8.5，如果(X,X)是演化稳定的，那么正常生物的预期收益要大于变种的预期收益，即$\varepsilon<3/4$，由此得$\bar{\varepsilon}=3/4$。

现在假设正常生物选择行动Y，而变种则选择行动X，那么可知一个正常生物的期望收益为

$$0\cdot\varepsilon+1(1-\varepsilon)=1-\varepsilon$$

一个变种的期望收益为

$$3\cdot\varepsilon+0(1-\varepsilon)=3\varepsilon$$

由此可得$\bar{\varepsilon}=1/4$，即当$\varepsilon<1/4$时，(Y,Y)是演化稳定的。

另外，可知(X,X)和(Y,Y)都是纳什均衡，并且它们都是严格纳什均衡，因而给定X，Y并不是X的最优反应，同样，给定Y，X并不是Y的最优反应，所以定义8.6同样满足，即(X,X)和(Y,Y)是演化稳定的。

再考虑图8－2所示博弈。假设正常生物选择的行动为X，而变种选择的行动为Y，同样变种在种群中所占的比重为ε。

	X	Y
X	3,3	1,1
Y	1,1	1,1

图8－2　生物配对博弈矩阵(2)

一个正常生物的预期收益为

$$3(1-\varepsilon)+1\cdot\varepsilon=3-2\varepsilon$$

一个变种的预期收益为

$$(1-\varepsilon)+1\cdot\varepsilon=1$$

因而当 $\varepsilon<1$ 时,(X,X)是演化稳定的。如果正常生物选择的行动为 Y,而变种选择的行动为 X,那么(Y,Y)是否为演化稳定的?可知,正常生物的预期收益为

$$1\cdot\varepsilon+1\cdot(1-\varepsilon)=1$$

一个变种生物的预期收益为

$$3\varepsilon+1\cdot(1-\varepsilon)=1+2\varepsilon$$

显然 $1>1+2\varepsilon$ 不可能,因而并不存在一个 ε,使得(Y,Y)是演化稳定的。根据定义 8.6,可知(X,X)为严格纳什均衡,所以(X,X)是演化稳定的。虽然(Y,Y)是纳什均衡,但是它不满足定义 8.6 中的第二个条件,因为给定 Y,显然 X 是 Y 的最优反应,但是 $u(Y,X)=1>u(X,X)=1$ 并不成立,所以(Y,Y)不是演化稳定的。

【例 8.5】 两虎相争

两只老虎争夺下崽的窝,夺得窝将能够繁殖出 $v>0$ 的幼仔,否则无窝下崽。两只老虎可能采取攻击(A)和退缩(B)两种行动。如果采取攻击,而对手退缩,那么将获得下崽的窝;如果采取攻击,而对手也采用攻击,那么收益为 $v/2-c/2$,其中 c 表示两虎相斗可能造成的伤害;如果两虎都退缩,那么收益为 $v/2$。其博弈矩阵如图 8-3 所示。现在我们求两虎相争博弈的演化稳定行动。

	A	B
A	$(v-c)/2,(v-c)/2$	$v,0$
B	$0,v$	$v/2,v/2$

图 8-3 两虎相争博弈矩阵

假如 $v>c$,那么可知该博弈有唯一的纳什均衡(A,A),其是严格纳什均衡,所以(A,A)是演化稳定的。实际上,A 还是严格优策略。

假如 $v=c$,那么可知存在唯一对称纳什均衡——(A,A)。对于(A,A)而言,其并不是严格纳什均衡,不过 $u(A,B)>u(B,B)$,所以(A,A)是演化稳定的。在上述两个情况下,可知在虎种群中,如果正常老虎为退缩,变种老虎为进攻,那么变种老虎将会使得虎种群发生改变,即变种老虎会逐步变成多数,而退缩的正常老虎会消失掉,在生物学中,把这种现象称为侵入。实际上当 $v=c$ 时,(A,B)和(B,A)同样是纳什均衡,但它们在进化过程中不会被保留下来。演化

博弈可以起到减少均衡数目的作用。

假如 $v<c$，那么可知不存在对称的（纯策略）纳什均衡，所以该博弈不存在演化稳定行动。

(易) 习题 8.5　在一个生物种群中进行两两配对。每个生物有两个行动，对应着 1、2 的收益需求。如果两个生物有一样的需求，那么各获得它所需求的；如果两个生物的需求不一样，那么需求小的生物获得对手的需求，而需求大的生物获得 $d\delta$，d 为它的收益需求，而 $\delta < 1/2$。请你找出该博弈的演化稳定行为。假如该博弈的每一个生物有 n 个行动，并且对应着 1、2、…、n 的收益需求，其中 $\delta < 1/n$，那么演化稳定行为又是什么？

上面我们所介绍的演化博弈只考虑了纯策略的情况，因而演化稳定行动有可能是不存在的。解决这个问题的思路主要有两个：一是考虑混合策略（或在种群中可能同时存在着多种类型的行为），二是考虑非对称均衡。

在允许混合策略的情况下，就有了著名的演化稳定策略（evolutionarily stable strategy，简称 ESS）定义。

定义 8.7　演化稳定（混合）策略　在一个对称双人策略博弈中，每一个参与者的预期收益为 $V(\alpha,\alpha')$，α 为当事者的混合策略，α' 为对手的混合策略。如果一个混合策略 α^* 是一个演化稳定策略（ESS），那么满足：

(1) (α^*,α^*) 是一个纳什均衡；

(2) $V(\beta,\beta) < V(\alpha^*,\beta)$ 对每一个 $\beta \neq \alpha^*$ 成立，其中 β 是 α^* 的最优反应。

定义 8.6 和定义 8.7 非常相似，但却并不等价，原因在于定义 8.6 只涉及纯策略（即退化的混合策略），而在定义 8.7 中却需要考虑所有的混合策略。一个演化稳定行动不一定是一个演化稳定策略，尽管纯策略也是混合策略。不过，如果行动 a^* 是严格优策略，那么 a^* 是演化稳定行动就一定是演化稳定策略（因为严格优策略对所有混合策略而言同样是严格优的）。为了说明上述结论，我们看图 8－4 所示的博弈。

	X	Y	Z
X	2,2	1,2	1,2
Y	2,1	0,0	3,3
Z	2,1	3,3	0,0

图 8－4　演化稳定行动不是演化稳定策略

根据定义 8.6，容易证明 (X,X) 为演化稳定行动，但它是否也是演化稳定策略呢？关键在于能否找到一个混合策略 β 是 α^*（以概率为 1 选择 X）的最优反应，并且 $V(\beta,\beta) \geq V(\alpha^*,\beta)$。不妨设 $\beta = (p,q,1-p-q)$，显然 β 是 α^* 的最

优反应。其次，$V(\beta,\beta)=2p^2+pq+p(1-p-q)+2qp+0+3q(1-p-q)+2(1-p-q)p+3(1-p-q)q+0=2p^2+3pq+3(p+2q)(1-p-q)$，$V(\alpha^*,\beta)=1+p$，显然，既是 α^* 的最优反应、又使得 $V(\beta,\beta)\geq V(\alpha^*,\beta)$ 成立的 β 是存在的，例如 $\beta=(0,2/3,1/3)$ 就满足上述要求。因而，(X,X)虽然是演化稳定行动，但却不是演化稳定策略。

【例 8.6】 混合策略下的两虎相争

仍然考虑两虎相争博弈，但允许混合策略。根据前面的分析，已知当 $v>c$ 时，(A,A)是演化稳定行动，而且容易证明(A,A)同样是演化稳定策略，因为以概率 1 选择 A 是纯策略 A 的唯一最优反应。当 $v<c$ 时，可知并不存在演化稳定行动，那么是不是也不存在演化稳定策略呢？答案是否定的。当 $v<c$ 时，容易证明当 $\alpha^*=(v/c,1-v/c)$ 时，(α^*,α^*) 是纳什均衡，并且 α^* 是 α^* 唯一的最优反应，所以 (α^*,α^*) 也是演化稳定策略。实际上，(α^*,α^*) 也是 $v=c$ 时的演化稳定策略。

虽然引人混合策略，但演化稳定策略仍然有可能是不存在的，例如，图 8－5 所示的博弈就不存在演化稳定策略，其中 $r>0$。

	X	Y	Z
X	r,r	$-1,1$	$1,-1$
Y	$1,-1$	r,r	$-1,1$
Z	$-1,1$	$1,-1$	r,r

图 8－5 不存在演化稳定策略

(易) 习题 8.6 证明图 8－5 所示博弈不存在演化稳定策略。（提示：如果 $r=1$，该博弈就变为剪刀、石头、布）

迄今为止，我们考虑的情景都是对称的：每一个参与者都完全一样。也正因为如此，没有必要标注出参与者的身份，这也是我们在图 8－1 至图 8－5 中没有标注出参与者 1 和参与者 2 的原因。如果参与者是不同的（或者具有不同的角色），那么情景就可能不是对称的，而且不对称的情景更加一般。那么如何模型化不对称情景，并且给出相应的演化稳定要求呢？一个思路就是将不对称情景变换成对称情景，这样前面所有结论就可以自然地运用在不对称情景中。一个假设就是参与者继承的是一种条件策略，即每一个参与者会根据其所处的地位（立场、角色等）、对手的地位及特性来决定其行动，如果我们将参与者的行动集合看作是由条件策略所组成，那么就可以将不对称情景转化成对称情景。

我们仍然通过例子来加以说明。与以前一样，一个种群中随机地两两配对，其中参与者 1 为守窝者，参与者 2 为抢窝者，该博弈有可能是不对称的。我们把该

博弈称为原博弈 C。我们不妨这样来看待每一个生物，它可能是守窝者，也可能是抢窝者，因而每一个生物的策略就是依据它的角色来选择行动——条件策略。一个生物的条件策略就是一个二元组(σ_1,σ_2)，σ_1 表示如果它是守窝者时定义在行动集合 A_1 上的混合策略，而 σ_2 则表示如果它是抢窝者时定义在 A_2 上的混合策略。经过这样处理，原博弈 C 就变成了一个对称博弈 G，因为每一个生物的条件策略集合都是一样的。在对称博弈 G 中，如果(σ_1,σ_2)为 i 的条件策略，(τ_1,τ_2)为 j 的条件策略，并且每一个生物成为守窝者的概率为1/2，那么 i 的收益为

$$u((\sigma_1,\sigma_2),(\tau_1,\tau_2)) = \frac{1}{2}U_1(\sigma_1,\tau_2) + \frac{1}{2}U_2(\tau_1,\sigma_2)$$

即当 i 为参与者 1 时，其收益为 $U_1(\sigma_1,\tau_2)$；当 i 为参与者 2 时，其收益为 $U_2(\tau_1,\sigma_2)$。

定义 8.8　非对称的演化稳定策略　一个条件策略是演化稳定的，假如它是 G 中的演化稳定策略，即条件策略(σ_1,σ_2)是演化稳定的，如果满足下列条件：

(1) $((\sigma_1,\sigma_2),(\sigma_1,\sigma_2))$是 G 中的纳什均衡；

(2) $u((\tau_1,\tau_2),(\tau_1,\tau_2)) < u((\sigma_1,\sigma_2),(\tau_1,\tau_2))$对于每一个条件策略$(\tau_1,\tau_2)\neq(\sigma_1,\sigma_2)$成立，其中$(\tau_1,\tau_2)$是$(\sigma_1,\sigma_2)$的最优反应。

同样，如果$((\sigma_1,\sigma_2),(\sigma_1,\sigma_2))$是 G 中的严格纳什均衡，那么第二个条件就是多余的。实际上，定义 8.8 不过是定义 8.7 的一个推广而已。由于混合策略不可能构成严格纳什均衡（如果是，就意味着某个纯策略的预期收益会大于其他纯策略的预期收益，这就与混合策略要求相矛盾，见第二章混合策略的内容），这就有了下列结论。

命题 8.2　一个条件策略是演化稳定的，当且仅当它是原博弈 C 中的一个严格（因而是纯的）纳什均衡。

【例 8.7】　不对称的两虎相争

两虎争夺下崽的窝，其中一只老虎为守窝者，而另一只老虎为抢窝者，V 为守窝者守窝成功时的收益，v 为抢窝者抢窝成功时的收益，其他内容与前面一样，不再赘述。原博弈 C 如图 8－6 所示。

		抢窝者	
		A	B
守窝者	A	$(V-c)/2,(v-c)/2$	$V,0$
	B	$0,v$	$V/2,v/2$

图 8－6　不对称的两虎相争博弈矩阵

当 $V<c$ 和 $v<c$ 时，容易证明存在两个严格纳什均衡——(A,B)和(B,A)。根据命题8.8，可知条件策略(A,B)(即如果是守窝者，那么选择A；如果是抢窝者，那么选择B)和条件策略(B,A)(即如果是守窝者，那么选择B；如果是抢窝者，那么选择A)是演化稳定的。我们也可以将原博弈 C 转化为 G，然后利用定义8.8来证明上述结论。可知条件概率(A,B)是对称博弈 G 的一个纯策略，显然策略组合((A,B),(A,B))是对称博弈 G 的一个纳什均衡，并且还是一个严格纳什均衡，所以(A,B)就是对称博弈 G 的演化稳定策略，条件策略(B,A)同理。

㊐ 习题8.7 仍考虑例8.8。原博弈 C 除了存在纯纳什均衡以外，还存在一个混合纳什均衡。考虑对称博弈 G 中的混合条件策略，并给出一个演化稳定的混合条件策略。

正像我们前面所提到的，演化(或进化)本质上是一个动态过程，通过动态分析将能够克服静态分析中的局限性，如果读者感兴趣，可以进一步阅读相关文献。

通过前面的介绍，我们可以得出这样的认识：演化稳定的思想实际上可以看作是纳什均衡在参与者进化过程中的一个扩展。另一方面，它也给了我们一些有益的启示。在生物界(包括我们人类在内)，都不具备所谓的完全理性，无所不知、无所不能是不可能的，绝大多数情况下我们所作的决策都是从最简单的原则开始，然后通过不断的演化从而达到最终找出最优策略的目的。例如，市场对资源最优配置的实现就是一个演化过程，虽然没有人能够智胜市场，但并不妨碍市场功能的发挥，也不会阻止亿万富翁的产生。从小的地方来看，下象棋也是一个演化过程，根据纳什定理，下象棋一定存在纳什均衡，但由于策略空间太过巨大，因而没有人能够事先计算出所谓的纳什均衡策略(最优策略)，通常人们都是从一些简单的套路开始(例如跳马、当头炮)，然后在下棋的过程中妙招连连(即逐步发现最优策略)，并最终决出胜负。所以说，演化博弈从另一个角度为博弈论的基本思想提供了新的基础。

第三节 学习博弈

学习博弈的一个重要假定就是博弈的参与者是有限理性的，因而随着博弈的进行(即演化)，不同的参与者之间的互动产生了一个学习的动态过程，其核心在于不同类型的博弈在相应的学习过程中如何导致长期博弈收敛于某一个纳什均衡。在学习过程中，参与者可能强化已有的信念、或者模仿或更新信念等，从而变得越来越老练和理性。在这一节中，我们只打算介绍最简单的有限理性

学习模型,这些模型又被称为强化模型,对于学习模型更深入、更全面的介绍,读者可以参阅相关的文献。

强化学习实际上反映了一种简单的机制,即参与者选择每一个策略时带有某种倾向,这可能是前辈教导的结果,也可能是一种习惯,还可能是一种经验,甚至是偶然的。采取某个策略的强化倾向一定与他们所获得的收益与潜在渴望水平的差异相一致,收益与渴望差异越大(如果为正),那么强化倾向就越强。

为了简单,我们假设强化是正的,渴望水平不变并等于0。时间是离散的,即 $t=0,1,2,\cdots$。两个参与者1和2,在每一个时刻 t 进行一个有限的策略式博弈 $G=\{S,u\}$,其中 $u(\cdot)>0$,$S_1=\{s_{11},\cdots,s_{1n}\}$,$S_2=\{s_{21},\cdots,s_{2n}\}$。在这些假设下,Posch(1997)给出了下面的强化学习模型。

每一参与者 i 在每一时刻 t 存在一个选择倾向向量 $\theta_i(t)=(\theta_{i1}(t),\cdots,\theta_{in}(t))$,令 $\Theta_i(t)=\sum_{j=1}^{n}\theta_{ij}(t)>0$,那么根据选择倾向向量 $\theta_i(t)$ 可以导出一个相应的权重向量 $\sigma_i(t)=(\sigma_{i1}(t),\cdots,\sigma_{in}(t))$,其中

$$\sigma_{ij}=\frac{\theta_{ij}(t)}{\Theta_i(t)} \tag{8-2}$$

$\sigma_i(t)$ 实际上给出了参与者 i 在时刻 t 选择每一个纯策略的概率,因而 $\sigma_i(t)$ 可以看作是参与者 i 在时刻 t 所采取的混合策略。相应的,参与者 i 在时刻 t 真正选择的纯策略为 $s_i(t)$。那么,参与者 i 在时刻 t 获得的收益就等于

$$u_i(t)=u_i(s_1(t),s_2(t)) \tag{8-3}$$

另一方面,对每一个纯策略 $s_{ij}\in S_i$,我们规定一个所谓的指标函数(indicator function),即

$$I_{ij}(t)=\begin{cases}u_i(t) & s_i(t)=s_{ij}\\ 0 & \text{否则}\end{cases} \tag{8-4}$$

我们不妨规定每一个参与者的选择倾向的运动规则为

$$\theta_{ij}(t+1)=\theta_{ij}(t)+I_{ij}(t) \qquad (i=1,2;j=1,\cdots,n) \tag{8-5}$$

这个运动规则体现出,强化可以看作是得到的收益与渴望水平之间的差异(正相关),由于渴望水平为零,所以参与者 i 所采取的每一个策略得到的强化就等于所得到的收益。

强化学习模型的关键就在于弄清楚每一个参与者 i 在任一时刻 t 由式(8-2)和式(8-5)规定的混合策略 $\sigma_i(t)$ 的调整过程,$i=1,2$。理想的调整过程应当是逐步收敛到一个稳定状态(比如纳什均衡)的。对于每一个纯策略 s_{ij},其概率的调整等于

$$\sigma_{ij}(t+1) = \frac{\theta_{ij}(t+1)}{\Theta_i(t+1)} = \frac{\theta_{ij}(t) + I_{ij}(t)}{\Theta_i(t) + u_i(t)}$$
$$= \frac{\theta_{ij}(t)}{\Theta_i(t)} + \frac{I_{ij}(t)\Theta_i(t) - u_i(t)\theta_{ij}(t)}{\Theta_i(t)[\Theta_i(t) + u_i(t)]}$$
$$= \sigma_{ij}(t) + \frac{I_{ij}(t) - u_i(t)\sigma_{ij}(t)}{\Theta_i(t) + u_i(t)} \quad (8-6)$$

需要指出的是,$I_{ij}(t)$和$u_i(t)$是随机的,它们取决于两个参与者在时刻t真正选择的具体纯策略。因而,对式(8-6)取数学期望,得

$$E_t[\sigma_{ij}(t+1) - \sigma_{ij}(t)] = \frac{E_t[I_{ij}(t) - u_i(t)\sigma_{ij}(t)]}{\Theta_i(t) + u_i(t)} \quad (8-7)$$

由于我们假设倾向和收益都为正,因而当$t\to\infty$时,$\Theta_i(t)\to\infty$;同时,$u_i(t)$和$I_{ij}(t)$有界,这就意味着式(8-7)的期望值在概率上趋于0,即随着时间的推移,混合策略在期望值上会收敛于一个稳定状态。实际上,经过简单的代数变换,式(8-7)将清楚地刻画出该系统的动态调整过程,这个动态调整过程在演化博弈中被称为复制者动态(replicator dynamics)。

Posch(1997)在其论文中,用一个2×2的策略式博弈对上述想法进行了证明,得出了如下的结论:

(1) 随机学习动态几乎一定(就是概率为1)收敛于一个封闭的轨迹(因为变量是时间的函数),即一个稳定点或是一个循环轨迹。

(2) 如果博弈存在严格均衡,那么学习动态一定收敛于某个严格均衡。

(3) 如果博弈不存在严格均衡,那么学习动态则展示出一个渐进的、连续的循环轨迹。

以上(1)实际上就是表明稳定状态是存在的,而(2)、(3)则是对(1)的具体说明。为了更好地理解(2)和(3)的含义,考虑如图8-7所给出的2×2策略式博弈。一般地,2×2博弈可以分成下面三类。

		参与者2	
		s_{21}	s_{22}
参与者1	s_{11}	a_{11},b_{11}	a_{12},b_{12}
	s_{12}	a_{21},b_{21}	a_{22},b_{22}

图8-7 一般的2×2策略式博弈

第一,如果$(a_{11}-a_{21})(a_{12}-a_{22})>0$,或者$(b_{11}-b_{12})(b_{21}-b_{22})>0$,即至少有一个参与者存在严格优策略,则只存在一个严格均衡,例如信用困境。

第二,如果$(a_{11}-a_{21})(a_{12}-a_{22})<0$,$(b_{11}-b_{12})(b_{21}-b_{22})<0$,$(a_{11}-a_{21})\cdot(b_{11}-b_{12})>0$,则存在两个(严格)纯策略均衡和一个混合策略均衡,例如性别战。

第三，如果$(a_{11}-a_{21})(a_{12}-a_{22})>0$，$(b_{11}-b_{12})(b_{21}-b_{22})<0$，$(a_{11}-a_{21})\cdot(b_{11}-b_{12})<0$，则只存在一个混合策略均衡，例如猜币游戏。

对于第一类的博弈，根据强化学习博弈，动态学习的结果是导致唯一的稳定预期——严格优均衡。例如，在讲授博弈论课程的过程中，让不同的同学做囚徒困境游戏，最初并非所有人都会选择坦白，但随着对博弈的学习，最后几乎所有同学都会选择坦白，即随着动态学习得出的结果只能是严格优均衡——唯一稳定点。

对于第二类的博弈，动态学习的结果是收敛到两个严格纯策略均衡，事实上，两个严格均衡出现的概率都大于0，各概率的大小则取决于收益和初始条件（历史、习惯）。至于第三个均衡——混合均衡，却是不稳定的。换言之，对混合均衡的任何小的扰动将导致偏离，而到达两个严格（纯策略）均衡中的任何一个。这意味着，动态学习实际上排除了混合均衡成为长期稳定状态的可能。实际上，恋爱中的情侣几乎不可能采用混合策略来决定约会地点，约会地点总是非常稳定的，即所谓的"老地方"。

对于第三类的博弈，其存在唯一的混合均衡，由强化学习动态导致的长期路径渐近地沿着其中一个封闭的轨迹围绕着纳什均衡运动。例如，在猜币游戏中，可能一开始参与者会始终选择正面，但随着博弈的进行，对手会逐步掌握这一点（学习），从而调整自己的行动，以至于最后的结果是1/2的时间选择正面，1/2的时间选择背面。

上面介绍的强化学习模型是最简单的，在它的基础上发展出了一系列新的学习模型，感兴趣的读者可以参阅有关文献。

我们在上面对博弈论及其纳什均衡的新发展和运用作出简单的介绍，从这些介绍可以看出，博弈论的运用领域是非常广阔的；同时，通过对演化、动态学习的研究，我们对博弈论中的各种均衡概念有了更深的理解，既为均衡提供了新的基础，又为求解均衡、精练均衡提供了简单有效的方法。

关键词

同盟博弈　大同盟　转移支付同盟博弈　具有凝聚力的同盟博弈　核　改善　单态纯策略模型　演化稳定行动　演化稳定策略　非对称的演化稳定策略　学习博弈

内容要点

1. 同盟博弈所要研究的就是组成利益集团(所谓的同盟)的利益所在,如果同盟不能给成员带来大于单打独斗的利益,同盟就难以维持,中国人所说的“树倒猢狲散”就是这个意思。同盟博弈最好的运用前景可能就是分析政治、军事和经济中的各种利益同盟。

2. 同盟博弈中的稳定概念与纳什均衡的思想本质上是一样的,分析同盟博弈的核是重中之重。所谓的核就是同盟博弈中稳定的同盟行动集合。命题8.1给出了检验核的行之有效的方法,有些时候可能它要烦琐一些,不过我们完全可以借助计算机程序来完成分析。

3. 演化博弈可以说从演化的角度对博弈论的一些基本概念和思想提供了来自于实践的支撑,从一些简单的原则出发,通过演化(竞争)最终实现最优。同时,博弈论运用于演化也为我们理解演化、乃至社会变革提供了强有力的工具。演化博弈中的均衡概念——演化稳定策略(行动)实际上可以看作是纳什均衡的一种精炼。

4. 学习博弈认为参与者都是有限理性的,通过学习不断地成熟,在互动学习过程中达到某个均衡。同时,通过学习博弈,还能够使我们剔除掉一些(不现实的)均衡,从而减少均衡数量,以达到精炼的目的。另外,学习博弈还可以这样来理解:它刻画了这样一个过程,参与者开始可以选择任一个策略,但随着不断学习的深入,博弈最终会收敛到一个均衡。

进一步阅读

Osborne(2004)对同盟博弈与核进行了通俗的说明,并提供了大量的案例分析。对演化博弈与学习博弈较正式和深入的讨论可参阅 Vega-Redondo(2003)。

参考文献

1. Abreu D. On the Theory of infinitely repeated games with discounting. Econometrica, 1988, 56: 383 – 396

2. Abreu D, Dutta P K and Smith L. The folk theorem for repeated games: a NEU condition. Econometrica, 1994, 62: 939 – 948

3. Barro R, Gordon D. Rules, discretion and reputation in a model of monetary policy. Journal of Monetary Economics, 1983, 12: 101 – 121

4. Benoit J P, Krishna V. Finitely repeated games. Econometrica, 1985, 53: 890 – 904

5. Benoit J P, Krishna V. Nash equilibria of finitely repeated games. International Journal of Game Theory, 1987, 16: 197 – 204

6. Bertrand J. (1883) Review by joseph bertrand of two books. Translated in Margaret Chevailier, History of Political Economy, 1992, 24: 646 – 653

7. Cho I K, Kreps D. Signaling games and stable equilibria. Quarterly Journal of Economics, 1987, 102: 179 – 222

8. Elster J. Marxism, functionalism and game theory. Theory and Society Ⅱ, 1982: 453 – 482

9. Feddersen T J, Pesendorfer W. Convicting the innocent: the inferiority of unanimous jury verdicts under strategic voting. American Political Science Review, 1998, 92: 23 – 35

10. Friedman J. A non-cooperative equilibrium for supergames. Review of Economic Studies, 1971, 38: 1 – 12

11. Fudenberg D, Maskin E S. The folk theorem in repeated games with discounting or with incomplete information. Econometrica, 1986, 54: 533 – 554

12. Fudenberg D, Tirole J. Game theory. Cambridge: MIT Press, 1991

13. Hardin R. Collective action as an agreeable n-prisoner's dilemma. Behavioral Science No. 5, 1971: 472 – 479

14. Harsanyi J. Games with imcomplete information played by Bayesian players, Part Ⅰ, Ⅱ. Management Science, 1967, 14: 159 – 182, 320 – 334

15. Harsanyi J. Games with imcomplete information played by Bayesian players,

Part Ⅲ. Management Science,1967,14:468 – 502

16. Hotelling H. Stability in competition. Economic Journal,1929,39:41 – 57

17. Hume D. (1739) A treatise of human nature. London: J. M. Dent, 1952 (Reprinted).

18. Gibbons R. Game theory of applied economists. Princeton: Princeton University Press,1992

19. Kreps D M. A course in microeconomic theory. Princeton: Princeton University Press,1990

20. Kreps D, Wilson R. Reputation and imperfect information. Journal of Economic Theory,1982,27:253 – 279

21. Mas-Colell A, Whinston M and Green J. Microeconomic theory. Oxford: Oxford University Press,1995

22. Milgrom P, Weber R J. A theory of auctions and competitive bidding. Econometrica,1982,50:1089 – 1122

23. Myerson R B. Incentive compatibility and the bargaining problem. Econometrica,1979,47:61 – 73

24. Myerson R B. Optimal auction design. Mathematics of Operations Research, 1981,6:58 – 73

25. Myerson R B. Bayesian equilibrium and incentive compatibility: an introduction. In Social Goals and social Organization. L Hurwicz, D Schmeidler and H Sonnenschein, eds. Cambridge: Cambridge University Press,1985

26. Myerson R B, and Satterthwaite. Efficient mechanisms for bilateral trading. Journal of Economic Theory,1983,28:265 – 281

27. Nash J. Equilibrium points in n-person games. Proceeding of the National Academy of Sciences of the United States of America,1950,36:48 – 49

28. Nash J. Non-cooperative games. Annals of Mathematics,1951,54:286 – 295

29. Nalebuff B. Credible pretrial negotiation. Rand Journal of Economics,1987, 18:198 – 210

30. Osborne M J. An introduction to game theory. Oxford: Oxford University Press,2004

31. Osborne M J, Rubinstein A. A course in game theory. Cambridge: MIT Press,1994

32. Rabin M. Psychology and economics. Journal of Economic Literature,1998, 36:11 – 46

33. Rawls J. A theory of justice, Ⅷ. Cambridge: Harvard University Press, 1971

34. Selten R. Spieltheoretische behandlung eines oligopolmodells mit nachfrage-tragheit. Zeitschrift fur Gesamte Staatswissenschaft, 1965, 121: 301 – 324

35. Selten R. Re-examination of the perfectness concept for equilibrium points in extensive games. International Journal of Game, 1975, 4: 25 – 55

36. Spence M. Job market signalling. Quarterly Journal of Economics, 1973, 87: 355 – 374

37. Taylor M. Anarchy and cooperation. London: Wiley, 1976

38. Thompson F B. (1952) Equivalence of games in extensive form. Research Memorandum 759, California: The Rand Corporation. In Classics in Game Theory. H. Kuhn, eds. Princeton: Princeton University Press, 1997: 36 – 45

39. Vega-Vedondo F. Economics and the theory of games. Cambridge: Cambridge University Press, 2003

40. Wen, Quan. The folk theorem for repeated games with complete information. Econometrica, 1994, 62: 949 – 954

41. Whitehead A N. Science and the modern world, New York: Mentor, 1948

42. 罗伯特·吉本斯. 博弈论基础. 北京: 中国社会科学出版社, 1999

43. 马丁 J. 奥斯本, 阿里尔·鲁宾斯坦. 博弈论教程. 北京: 中国社会科学出版社, 2000

44. 迈尔森. 博弈论——矛盾冲突分析. 北京: 中国经济出版社, 2001

45. 奥尔森. 集体行动的逻辑. 上海: 上海三联书店, 1990

46. R. J. 奥曼. 博弈论. 新帕尔格雷夫经济学大辞典Ⅱ. 北京: 经济科学出版社, 1996.

47. 费尔南多·维加－雷东多. 经济学与博弈理论. 上海: 上海人民出版社, 2006

48. 张维迎. 博弈论与信息经济学. 上海: 上海三联书店, 上海人民出版社, 1996

人名索引

Abreu,160,174
Barro,171
Benoit162,174
Bertrand,78
Cho255,268
Elster,13
Feddersen,197
Friedman,158,174
Fudenberg,77,160,174
Gibbons,106,141
Gordon,171
Green,77,268
Hardin,15
Harsanyi,217
Hotelling,81
Hume,87
Kreps,25,77,141,218,244,255,268
Maskin,160,174
Mas-Colell,77,268
Milgrom,207
Myerson,207,211,212,217
Nalebuff,249,254
Nash,26,72,77,95,106
Osborne,106,141,217,268
Pesendorfer,197
Rabin,25
Rawls,15
Rubinstein,141,217
Satterthwaite,211
Selten,117
Smith,160
Spence,12,235,259,268
Taylor,15
Thompson,115
Tirole,77
Vega-Vedondo,284
Weber,207
Wen,174
Whinston,77,268
Whitehead,13,192
奥尔森,13,192
奥曼,25
张维迎,25,106

主题索引

v－N－M收益函数,63
阿基米得公理,63
半混同均衡,237
贝叶斯博弈的基本式,182
伯努利(Bernouli)收益函数,63
伯努利收益函数,144
博彩,219
博弈,4
博弈规则,16
博弈论,6
不动点,73
不可置信的威胁,116
布劳尔(Brower)不动点定理,73
参与者,16
参与者函数,109
策略,114
策略博弈中的演化稳定行动,274
策略的相互依存性,2
策略式,115
触发策略,150
纯策略,61
纯策略贝叶斯纳什均衡,185
纯公共价值,205
纯过错,84
搭便车行为,13
单边背离特性,152
单调变换,9
单态纯策略模型,273
道德风险,12
等价性定理,121
等收益定理,199
动态贝叶斯博弈的扩展式,220
动态博弈,22,115
动态不一致的,116
动态一致性,107
对称博弈,59
对称纳什均衡,60
二级密封式价格拍卖,198
泛子博弈,224
非对称的演化稳定策略,280
非完全信息博弈,21
非严格均衡,47
分离均衡,237
改善,271
公共价值,198
公共价值等收益定理,207
公共信息,21
公共知识,22
过错加过错,84
豪尔绍尼转换,177
荷式拍卖,198
混合策略,62
混合策略贝叶斯纳什均衡,187
混合策略纳什均衡,66
混合策略下,71
混合策略下的对称博弈,71
混同均衡,237
基本式,28





激励兼容,213
极小极大收益,159
角谷(Kakutani)不动点定理,74
阶段博弈,144
结果,16
结果函数,114
结果路径,145
紧集,75
静态博弈,22
具有凝聚力的同盟博弈,271
决策主体,11
绝对责任,84
均衡,11
可行收益组合,157
空历史,109
扩展式博弈,115
类型,184
理性的,9,63
理性选择原理,9
连锁店悖论,244
连续策略下,104
劣策略,37
纳什定理,72
纳什均衡,39,51
纳什均衡的等价性,66
逆推法,118
逆向选择,12
偏好,8
全历史,109
上半连续,74
审时度势,226
市场失灵,12
收益,16
收益函数,9
双人谈判问题,93
私人价值,198
谈判问题的解,94
替代公理,63
贴现平均收益值,143
贴现因子,143
同盟博弈,269
同盟博弈的核,271
凸集,75
完美贝叶斯均衡,226
完美信息动态博弈,109
完全信息博弈,21
完全信息动态博弈的扩展式,110
无名氏定理,158
无限重复2人无名氏定理,160
无限重复n人无名氏定理,160
无效率定理,211
显示原理,213
信号,181
信号博弈,235
信号函数,181
信念,182
信念系统,225
信念与策略的一致性,229
信息分割,219
信息集,219
行为,11
行为策略,225
序列均衡,229
序列均衡的再精炼,255
序列理性,229
学习博弈,280
严格均衡,47
严格劣策略,36,65

严格优策略,34
严格优策略均衡,34
演化稳定(混合)策略,277
演化稳定的,274
一报还一报策略,154
一个子博弈,117
一级密封式价格拍卖,198
隐集团,14
英式拍卖,197
赢者的诅咒,206
优策略,35
优势准则,255
有限惩罚策略,153
有限的,72,120
有限重复2人无名氏定理,162
有限重复 n 人无名氏定理,162
杂合均衡,237
真子历史,109
直观准则,258
直接机制,213
制度结构,11
重复博弈,144
重复博弈的扩展式,145
重复直观准则,259
转移支付同盟博弈,270
状态,181
状态类,181
子博弈完美均衡,118
子历史,109
自然,218
最优反应函数,50